# Food subsidies in general equilibrium

Hummel, A.J.; Ziesemer, Vinzenz

**Publication date**
2021

CESifo WORKING PAPERS

9201
2021

July 2021

# Food Subsidies in General Equilibrium

*Albert Jan Hummel, Vinzenz Ziesemer*

CESifo

**Impressum:**

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9201*

# Food Subsidies in General Equilibrium


## 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. 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 food subsidies in the range 5%-12%.


JEL-Codes: E640, H210, Q180.

Keywords: uniform consumption taxes, general equilibrium effects, food subsidies.

*Albert Jan Hummel*
*University of Amsterdam*
*Amsterdam / The Netherlands*
*a.j.hummel@uva.nl*

*Vinzenz Ziesemer*
*Erasmus University Rotterdam*
*Rotterdam / The Netherlands*
*ziesemer@euc.eur.nl*

July 16, 2021
We would like to thank 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.

# 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 increase 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-skill 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 more important.

We study the optimal design of food subsidies both theoretically and quantitatively. We follow the setup of Deaton (1979), in which a government redistributes using linear income and consumption taxes. This is the instrument set we are most interested in.[3] 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).[4] 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. Low-income households spend a larger fraction of their income on food, but Engel curves are still linear. Consequently, optimal food subsidies are zero in the absence of general equilibrium effects (Deaton, 1979).

Theoretically, we show that standard formulas for optimal linear taxes (cf. Sheshinski (1972), Feldstein (1973) and Diamond (1975)) are modified if taxes induce general equilibrium effects on prices. We derive an expression for the optimal food subsidy in terms of sufficient statistics, in the spirit of Saez (2001) and Chetty (2009). The additional statistics that are relevant for determining the welfare effect of a change in the food subsidy concern the effect of the subsidy on equilibrium prices – in particular the wages for high-skilled and low-skilled workers. Changes in equilibrium prices have both distributional and fiscal effects, which the government should take into account when designing tax policy.

We then proceed to investigate quantitatively the implications of general equilibrium effects for the optimal design of food subsidies by calibrating the 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 expenditure 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 use the 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

[3] In an extension, we allow for nonlinear taxes on labor income.

[4] 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 extensively in the concluding section of this paper.

policies, including a food subsidy rate of 10% (OECD, 2017), complete our parameterization.

Turning to the results, we find that the optimal food subsidy is close to 5% if the government has a utilitarian objective and increases to approximately 12% if the government has a Rawlsian (maxi-min) objective. Hence, general equilibrium effects rationalize fairly sizable deviations from uniform consumption taxes. Despite this, the welfare gains of using food subsidies are rather modest. Starting from a setting with uniform consumption taxes, a utilitarian government is indifferent between optimizing food subsidies and increasing everyone's consumption aggregate by 0.01%. This figure increases to 0.07% if the government has a Rawlsian objective. We also show that the welfare effects are much larger if food subsidies are severely mis-optimized. For example, if food is subsidized at a rate of 60%, the welfare gains from optimizing them can be as large as 3% in consumption equivalents.

Given the considerable attention paid to consumption taxes in the literature and in policy discussions, we consider these 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 are small, unless current subsidy rates are very large.

We investigate the robustness of our main results by varying the elasticity of substitution i) between skill types in the production function and ii) between agricultural and non-agricultural goods in the utility function. The first of these determines the strength of general equilibrium effects, whereas the second determines by how much individuals change their consumption mix in response to a change in the food subsidy. All in all, we find that both optimal food subsidies and the welfare gains are fairly robust to variations in these parameters. We also study an extension where labor income can be taxed nonlinearly. This takes away some of the rationale for the food subsidy, as a nonlinear income tax allows the planner to employ general equilibrium effects even before the use of food subsidies. Consequently, we find that allowing for nonlinear income taxes reduces both the optimal food subsidy and the corresponding welfare gains, but qualitatively our results remain intact.

Finally, we study a number of reforms to the current tax system. Our results highlight that general equilibrium effects are of the same magnitude as the more standard effects that show up in optimal tax formulas (see Saez, 2001 and Chetty, 2009). Consequently, ignoring general equilibrium effects can lead policymakers astray when judging the desirability of tax reforms. We show that this is the case if a Rawlsian government considers reforming the current food subsidy.

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 derives an expression for the optimal food subsidy. Section 4 discusses the data and the parameterization. Section 5 contains the quantitative analysis of optimal taxes, along with several robustness checks. Section 6 analyzes tax reforms. Section 7 concludes and discusses directions for future research.

## 1.1 Related literature

How to optimize consumption taxes is an age-old problem in public economics. Ramsey (1927) shows that the optimal tax rate 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 weak separability between consumption and leisure in the utility function is enough to render consumption taxes redundant provided the government can levy a nonlinear tax on labor income. Deaton (1979) analyzes a model where the government levies a linear tax on labor income. He finds that the result by Atkinson and Stiglitz (1976) stands as long as Engel curves are linear. In both these 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), though we consider nonlinear income taxes in an extension. The main 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 Engel curves are linear and preferences are weakly separable between consumption and leisure.

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) studies 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). The optimal tax rules differ from those from the classic analysis by Mirrlees (1971) because changes in wages alleviate or tighten incentive constraints. Naito (1999) extends the analysis from Stiglitz (1982) by including different consumption goods. 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. This is the case if consumption taxes compress the (pre-tax) wage distribution.

Our paper differs from these analyses, in particular those of Stiglitz (1982) and Naito (1999), in two substantive ways. First, in most of what follows we consider a setting with linear tax instruments similar to that of Allen (1982), who abstracts from consumption taxes. This allows us to derive optimal tax formulas in terms of sufficient statistics, in the spirit of Saez (2001) and Chetty (2009). Second, and more importantly, we attempt to quantify the optimal tax or subsidy on agricultural products and 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. (2020) and Doligalski and Rojas (2020) consider the effect of the informal sector on optimal taxation. Bachas et al. (2020) study policy implications for food subsidies, but do not take into account the general equilibrium effects that we investigate in this paper. Our paper's focus on the importance of general equilibrium effects in consumption taxation complements this literature.

# 2 Model

In this section, we describe the model that is used in the remainder of the analysis. The economy contains two types of individuals, low-skilled $L$ and high-skilled $H$. Each type is present with mass $\mu_i$, where $i \in \{L, H\}$ and the total mass equals one: $\mu_L + \mu_H = 1$. Production takes place in two sectors $j \in \{a, n\}$, denoting agriculture and non-agriculture, in which returns to scale are constant.[5] 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 $w_i$ that does not depend on the sector of employment. Finally, there is a government that has a preference for redistribution. It 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.

[5]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 $\ell$. Given taxes and given prices, an individual of type $i \in \{L, H\}$ solves the following maximization problem:

$$\max_{\{c_{a,i}, c_{n,i}, \ell_i\}} \quad V_i = u(c_{a,i}, c_{n,i}) + v(\ell_i)$$
$$\text{s.t.} \quad p_a(1+\tau_a)c_{a,i} + c_{n,i} = T + w_i(1-\tau_y)\ell_i. \tag{1}$$

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 $\sigma$ and $\phi$ regulate the curvature of the consumption aggregate and labor supply in utility, while $\psi$ governs their relative importance. The relative importance of agricultural goods in consumption is regulated by $\omega$, while $\epsilon$ 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. 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). Hence, any departure from uniform consumption taxation is 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 wage $w_i$ is indexed by skill type. By contrast, the tax instruments are the same for all individuals and do not depend on, say, skill or sector of employment. Tax policy consists of a linear tax rate $\tau_a$ on agricultural goods and a linear tax rate $\tau_y$ 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. The reason is that 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. Therefore, we 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:

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

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

where $P$ is a price aggregator, or index:

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

The first of these equates the marginal rate of substitution between leisure and the consumption aggregate to the relative price. 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:

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

Here, total factor productivity $A_j$ and the share parameter $\gamma_j$ 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 $\rho$ between the two skill types is assumed to be the same in both sectors.

A representative firm in each sector $j \in \{a, n\}$ maximizes profits by choosing how much labor of each skill type to hire, 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: $\Pi_j = 0$. 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, the wages $w_L$ and $w_H$ and hence the skill premium do not vary across 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 $\gamma_i$ is highest).

## 2.3 Government

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

$$\mathcal{W} = \mu_L \alpha_L V_L + \mu_H \alpha_H V_H. \tag{12}$$

Here, $\alpha_L \geq 0$ and $\alpha_H \geq 0$ are exogenous Pareto weights, which, together with concavity in the individual utility function, determine the government's preferences for redistribution. The sum of these weights is normalized to one: $\alpha_L + \alpha_H = 1$. A utilitarian objective corresponds to setting $\alpha_L = 1/2$ and a Rawlsian (maxi-min) objective corresponds to setting $\alpha_L = 1$.

The government consumes an exogenous amount of $G$ units of non-agricultural goods. Consequently, its budget constraint is given by

$$\tau_a \sum_i \mu_i p_a c_{a,i} + \tau_y \sum_i \mu_i w_i \ell_i = T + G, \tag{13}$$

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 optimal tax problem.

## 2.4 Equilibrium

A competitive equilibrium is formally defined as follows:

**Definition 1.** *A competitive equilibrium consists of consumption and labor supply decisions* $\{(c_{a,L}, c_{n,L}, \ell_L), (c_{a,H}, c_{n,H}, \ell_H)\}$*, labor inputs* $\{(L_a, H_a), (L_n, H_n)\}$ *and prices* $(p_a, w_L, w_H)$ *such that, given tax policy* $(\tau_a, \tau_y, T)$*,* *(i) individuals maximize utility, cf. equations (1), (4) and (5),*

*(ii) firms maximize profits, cf. equations (9)–(10),*

*(iii) labor and goods markets clear:*

$$L_a + L_n = \mu_L \ell_L, \tag{14}$$

$$H_a + H_n = \mu_H \ell_H, \tag{15}$$

$$Y_a = \mu_L c_{L,a} + \mu_H c_{H,a}, \tag{16}$$

$$Y_n = \mu_L c_{L,n} + \mu_H c_{H,n} + G. \tag{17}$$

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 $(\tau_a, \tau_y, T)$. 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 satisfied.

# 3 Optimal food subsidies

The government chooses tax policy $(\tau_a, \tau_y, T)$ 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 depend on the tax instruments. The optimal tax problem is formally analyzed in Appendix A. Because our focus is on food subsidies, we only state the main result regarding the optimal $\tau_a$ here.

**Proposition 1.** *At the optimal tax system, the following condition must hold:*

$$\frac{\partial \mathcal{W}}{\partial \tau_a}\frac{1}{\lambda} = DE + BE + GE = 0, \tag{18}$$

*where $\lambda$ is the multiplier on the government budget constraint, $DE$ stands for 'direct effects', $BE$ for 'behavioral effects' and $GE$ for 'general equilibrium effects'. These are given by*

$$DE = \sum_i \mu_i (1 - g_i) p_a c_{a,i}, \tag{19}$$

$$BE = \tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial \tau_a} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial \tau_a}, \tag{20}$$

$$GE = \sum_i \mu_i \Big[\tau_a - g_i(1+\tau_a)\Big] c_{a,i} \frac{\partial p_a}{\partial \tau_a} + \sum_i \mu_i \Big[\tau_y + g_i(1-\tau_y)\Big] \ell_i \frac{\partial w_i}{\partial \tau_a}. \tag{21}$$

*Here, the summation takes place over skill types $i \in \{L, H\}$ and $g_i = \frac{\alpha_i u_{n,i}}{\lambda}$ denotes the welfare weight, which measures the increase in social welfare if an individual of type $i$ receives an additional unit of income.*

*Proof.* See Appendix A. □

A change in $\tau_a$ has three welfare-relevant effects, which at the optimum sum to zero. First, a higher tax on agricultural goods transfers income from individuals to the government budget. We label these 'direct effects'. How much is transferred from individuals of skill type $i \in \{L, H\}$ to the government budget depends on their population share $\mu_i$ and their spending on agricultural goods $p_a c_{a,i}$. The corresponding welfare effect is proportional to $1 - g_i$, which measures the benefit of transferring one unit of income from an individual of type $i$ (whose welfare weight is $g_i$) to the government (whose weight is one). The term $DE$ sums these effects over the two skill types.

Second, a change in the tax on agricultural products $\tau_a$ 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. If individuals were 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 consumption are taxed or subsidized. These are the fiscal externalities that individuals do not take into account when making their consumption and labor supply decisions. Naturally, the welfare-relevant effects due to these fiscal externalities are proportional to $\tau_a$ and $\tau_y$, respectively. The term $BE$ adds these fiscal effects over the individual skill types and over the two tax bases.

Third, a higher tax on agricultural goods affects equilibrium prices $p_a$, $w_L$ and $w_H$. We label these 'general equilibrium effects'. To illustrate, suppose that a higher tax $\tau_a$ 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 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 generates a fiscal externality that is proportional to $\tau_a$, whereas changes in equilibrium wages $w_L$ and $w_H$ generate a fiscal externality that is proportional to $\tau_y$. Second, changes in equilibrium prices have an effect on individual utility: a higher price of agricultural goods lowers resources available for consumption, whereas a higher wage has the opposite effect. These effects are weighed by the welfare weights $g_i$. The term $GE$ then sums the welfare-relevant effects over skill types and over the price responses.

At the optimum, the sum of the direct, behavioral and general equilibrium effects is equal to zero. Equation (18) gives an optimal tax formula that is expressed in terms of population shares, welfare weights and the responses of equilibrium quantities and prices to changes in $\tau_a$. These are the 'sufficient statistics' that determine the optimal tax on agricultural products goods (Chetty, 2009). Similar expressions can be derived for the optimal tax on labor income $\tau_y$ and the optimal transfer $T$: see Appendix A. It is worth pointing out that these formulas hold under a general specification of preferences and technology. In that sense, our choice for the utility and production function (combined with the parameterization discussed below) can be seen as a way of pinning down these statistics.

## 3.1 General equilibrium effects

Determining the sign of the optimal tax $\tau_a$ on agricultural products turns out to be a formidable task, even after having made specific functional form assumptions. Therefore, instead of attempting a formal proof, we conjecture (and verify numerically) that the optimal tax on agricultural goods is negative if agricultural production is relatively low-skilled labor intensive. Hence, we conjecture that the optimal $\tau_a < 0$ if $\gamma_a > \gamma_n$.

To understand why, first consider the instructive case where labor types are perfect substitutes and none of the skill types has a comparative advantage in the production of agricultural goods: $\rho \to \infty$ and $\gamma_a = \gamma_n = \gamma < 1/2$. Equilibrium wages are then given by $w_L = \gamma$, $w_H = (1 - \gamma)$ and the skill premium is $w_H/w_L = (1 - \gamma)/\gamma > 1$. The price of agricultural goods, in turn, equals $p_a = 1/A_a$. None of these prices depends on any of the tax instruments. Hence, there are no general equilibrium effects associated with an increase in the tax on agricultural goods: $GE = 0$. Because preferences are separable between consumption and leisure and Engel curves are linear, we know from Deaton (1979) that all goods should be taxed at the same rate: the optimal $\tau_a = 0$. It follows immediately that under the current specification of preferences, any departure from uniform consumption taxation (i.e., an optimal $\tau_a$ that differs from zero) must be driven by general equilibrium effects.

To allow for general equilibrium effects, suppose labor types are imperfect substitutes in production and low-skilled workers have a comparative advantage in the production of food: $\rho \ll \infty$ and $\gamma_a > \gamma_n$. From equation (11) it is clear that the production of agricultural goods is relatively low-skilled labor intensive: $L_a/H_a > L_n/H_n$. In this case, there are general equilibrium effects that can be exploited for redistributive purposes.

To illustrate, suppose the government lowers the tax (or raises the subsidy) on agricultural products: $\tau_a$ decreases. This leads to an increase in the demand for agricultural products, which in turn raises the demand for low-skilled labor. A higher demand for low-skilled labor reduces the skill premium $w_H/w_L$. This is essentially an application of the Stolper-Samuelson theorem, with the increase in demand for a particular good driven by a reduction in the tax rather than an opening to trade. A lower skill premium indirectly redistributes income from high-skilled to low-skilled workers, which *ceteris paribus* raises welfare.

Starting from a situation with uniform consumption taxation (i.e., starting from $\tau_a = 0$), the distributional benefits driven by general equilibrium effects provide a rationale for introducing a food subsidy. Doing so, however, comes at the costs of distorting consumption decisions. After the imposition of a subsidy, consumers face a price of agricultural goods in terms of non-agricultural goods that differs from the marginal rate of transformation. This leads to inefficiently high levels of food consumption, as individuals do not internalize the negative impact of food consumption on the government budget. These costs are zero if $\tau_a = 0$ (i.e., starting from a situation with uniform consumption taxes), but increase as the food subsidy rises – typically at an increasing rate (Harberger, 1964). This ultimately puts a limit on the extent to which food should be subsidized.

The optimal food subsidy balances the indirect distributional benefits that come from general equilibrium effects on wages against the costs of distorting consumption decisions. How large optimal food subsidies are once we account for general equilibrium effects is the topic of the following sections.

# 4 Parameterization

We parameterize our model to represent the Chinese economy in the year 2008 as well as possible. We combine a variety of data sources, most importantly the 2008 sample of the Chinese Household Income Project (CHIP). Below, we begin by discussing the CHIP data. We then 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 was 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 facts 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 the majority of hours worked.[6] 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 results in a relative low-skill use of agriculture versus non-agriculture of 1.14.[7] Matching these three figures implies matching the entire two-by-two table. In addition, we target the college premium observed in the data. Our dataset contains few observations on wages by educational status, but Wang (2012) investigates the topic in depth and finds a (non-causal) college premium of 51% using data from CHIP 2002.

The CHIP data also includes information on consumption expenditure for the urban popu-

[6]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).

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

| | 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

lation. 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 $\omega$ and $\underline{c}_a$.[8] 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.

Unfortunately, the data from the expenditure section of the CHIP 2008 survey do not align 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.[9] We prefer the calibrated model to be consistent with macro aggregates, and therefore choose to shift downward the linear Engel curve to match an expenditure share in

[8]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 $\tau$, 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. 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 $\tau$, which we found to be 0.17, and of $S_H^u$, which we found to be 19.6%.

[9]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%.

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

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.[11]

We also require values for a number of elasticities. For many of these, no specific estimates are available for China. For this reason, we rely on the idea that behavioral is similar across countries. We set the coefficient of relative risk-aversion $\sigma$ equal to one (logarithmic utility). Moreover, we use the parameters $\phi$ and $\psi$ governing the disutility of labor to target an average Frisch elasticity of 0.5 and an average Hicksian elasticity of 0.33 (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

[10] 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.

[11] 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.

| Parameter | Value | Moment | Model | Data |
|---|---|---|---|---|
| $\mu_L$ | **0.90** | Share of low-skill workers | 90.2% | 90.2% |
| $\sigma$ | **1.01** | Coefficient of relative risk aversion | 1.01 | 1.01 |
| $\phi$ | 1.35 | Average Frisch elasticity | 0.50 | 0.50 |
| $\psi$ | 0.53 | Average Hicksian elasticity | 0.33 | 0.33 |
| $\omega$ | 0.39 | Slope of expenditure regression | 0.25 | 0.25 |
| $\epsilon$ | **0.50** | Consumption elasticity of substitution | 0.50 | 0.50 |
| $\underline{c}_a$ | 0.14 | Food share of expenditures | 26.2% | 26.1% |
| $A_a$ | 3.56 | Agriculture share of employment | 26.5% | 26.6% |
| $\gamma_a$ | 0.95 | Relative low-skill intensity across sectors | 1.14 | 1.14 |
| $\gamma_n$ | 0.72 | College premium | 1.51 | 1.51 |
| $\rho$ | **1.40** | Elasticity of substitution between types | 1.40 | 1.40 |
| $\tau_a$ | **-0.10** | Agricultural producer support as % of receipts | 10.0% | 10.0% |
| $\tau_y$ | **0.15** | Effective tax burden on labor income | 14.5% | 14.5% |
| $T$ | 0.01 | Tax revenue as % of GDP | 10.1% | 10.1% |

The table informally groups the data moments with the parameters they are considered informative of, although each parameter influences many moments. Of the 14 parameters, 6 are set directly to match their empirical counterpart. These 'outside' parameters are indicated in boldface.

Table 2: Parameters and moments

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 respect to this parameter choice as well.

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 Optimal tax analysis

This section presents the optimal tax analysis. It begins with our main results: general equilibrium effects rationalize sizable deviations from uniform consumption taxes. At the same time, the welfare gains that come from optimizing food subsidies are very modest. We then proceed to discuss the robustness of our findings by varying the elasticity of substitution i) between skill types in production ($\rho$) and ii) between agricultural and non-agricultural goods in the consumption aggregate ($\epsilon$). Our main findings are robust to reasonable variations in these parameters. Finally, we study optimal policies if the government can levy a nonlinear tax on labor income. This shrinks both the size and the welfare gains of optimal food subsidies, but leaves the results qualitatively intact.

## 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 using them.

Table 3 shows the optimal taxes along some statistics on the resulting allocation.[12] 'Baseline' refers to the calibrated model, which serves as a comparison to the other columns. The results for the utilitarian criterion are obtained by setting $\alpha_L = \alpha_H = 1/2$ and for the Rawlsian criterion by setting $\alpha_L = 1$ and $\alpha_H = 0$. The columns 'Optimal' show the results if no restrictions are imposed on food subsidies. By contrast, the results under 'Uniform' are obtained by imposing food subsidies are abandoned: $\tau_a = 0$. These columns plot the results if all goods are taxed at the same rate.

| | Baseline | Utilitarian | | Rawlsian | |
|---|---|---|---|---|---|
| | | Optimal | Uniform | Optimal | Uniform |
| Tax on agricultural goods ($\tau_a$) | –10.00% | –4.68% | 0.00% | –12.32% | 0.00% |
| Tax on labor income ($\tau_y$) | 14.53% | 7.28% | 6.43% | 18.04% | 16.07% |
| Transfer/average income ($T/\bar{y}$) | 1.90% | –3.77% | –3.46% | 4.67% | 5.85% |
| Skill premium | 1.51 | 1.53 | 1.53 | 1.50 | 1.52 |
| Agriculture share in GDP | 25.38% | 24.88% | 24.55% | 25.62% | 24.70% |
| Share of low-skilled in agriculture | 29.15% | 28.59% | 28.21% | 29.41% | 28.39% |
| Consumption equivalent gain | N/A | 0.01% | | 0.07% | |

Table 3: Optimal policy

Under the utilitarian criterion, the optimal food subsidy equals 4.68%. The preferential tax treatment of food relative to the untaxed numeraire (non-agricultural goods) is therefore

[12]These results are obtained by numerically solving the optimal tax problem as described in Appendix A. Additional details and Matlab codes are available upon request.

modest.[13] Furthermore, income is also taxed at a fairly low rate of 7.28%. The government budget constraint (13) then implies the lump-sum transfer is negative: all individuals pay a lump-sum tax that corresponds to approximately 3.77% of average income.[14] Perhaps surprisingly, the optimal tax system under a utilitarian criterion is less redistributive than the current system: compare columns two and three. This is no longer the case if the government has a Rawlsian objective and attempts to maximize the well-being of low-skilled workers. In this case, the optimal food subsidy is 12.32%. A Rawlsian planner thus uses larger food subsidies to boost the wage of low-skilled workers. The optimal tax rate on labor income, in turn, is 18.04% and individuals receive a lump-sum transfer that corresponds to 4.67% of average income. All these figures are larger than their counterparts in the baseline: compare columns two and five.

It appears that the current policy lies somewhat in between the one that would be chosen by a utilitarian and the one that would be chosen by a Rawlsian planner. This is confirmed in Section 6.3, where we show that, starting from the current system, the room for Pareto improvements is limited. Despite this, we want to stress that the reasons we observe food subsidies in practice are likely to be very different from the considerations discussed in the above. For example, governments may introduce these subsidies under pressure from interest groups. Lopez et al. (2017) study changes in China's agricultural subsidies, and find that larger and more geographically concentrated parts of the agricultural sector are subsidized at a higher rate. Whatever explains the set of policies currently in place, our results should not be read as an attempt to rationalize them.

The columns labeled 'Uniform' show 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 constrained optimal policies. Comparing the results with the unconstrained policies, we see that the optimal tax rate on labor income is lower under both welfare criteria. Intuitively, there is no need to finance food subsidies. This also explains why the optimal lump-sum tax is smaller under the utilitarian objective, and why the optimal lump-sum transfer is larger under the Rawlsian criterion.

The fourth, fifth and sixth row of Table 3 illustrate the mechanism through which food subsidies improve welfare. The first of these shows the skill premium under a utilitarian

[13] We verify numerically that optimal food subsidies are zero if low-skilled workers do not have a comparative advantage in the production of agricultural goods, i.e., if $\gamma_a = \gamma_n$. In this case, there are still general equilibrium effects in the sense that prices are endogenous to tax policy (as $\rho \ll \infty$), but food subsidies are ineffective at raising the wage of low-skilled workers relative to that of high-skilled workers.

[14] The lump-sum tax does not translate into negative consumption, as there are no individuals with zero income. The reason why the lump-sum transfer is negative is that the revenue requirement $G$ is fairly sizable. Without a revenue requirement, the lump-sum transfer is positive.

and Rawlsian objective, with ('Uniform') and without ('Optimal') imposing the restriction that $\tau_a = 0$. If a utilitarian government follows a uniform consumption tax policy, the skill premium is approximately 1.53. If the government can also optimize food subsidies, the skill premium decreases by 0.53 percentage points (although it stays at 1.53 when rounding to two digits). The decrease for a Rawlsian planner is somewhat more pronounced. Starting from the constrained optimum with uniform consumption taxes, optimizing food subsidies generates a 1.39 percentage point reduction in the skill premium (to 1.50).

Significant changes in the size of the agricultural sector drive these decreases in the skill premium, as illustrated by the fifth and sixth row of Table 3. These rows show the share of agriculture in GDP and the fraction of low-skilled workers that are employed in agriculture. Under a utilitarian objective, optimal food subsidies raise the agricultural share in GDP by 0.34 percentage points (to 24.88%) compared to the uniform case. Similarly, the fraction of low-skilled workers employed in agriculture increases by 0.37 percentage points (to 28.59%). As before, the increases are larger if the government has a Rawlsian objective. In that case, optimal food subsidies increase the agricultural share in GDP by 0.92 percentage points (to 25.62%) and the fraction of low-skilled workers employed in agriculture by 1.02 percentage points (to 29.41%).

The last row of Table 3 shows the welfare gains that can be obtained from using food subsidies or, equivalently, the welfare costs of following a uniform consumption tax policy. Specifically, starting from the constrained optimal policy with $\tau_a = 0$, the government is indifferent between optimizing food subsidies and increasing the consumption aggregate $C(c_{a,i}, c_{n,i})$ of all individuals by the reported consumption equivalent gain. As Table 3 shows, the welfare gains are modest. A utilitarian government is indifferent between using food subsidies and increasing everyone's consumption aggregate by 0.01%.[15] Not surprisingly, this figure is larger for a Rawlsian government, which also uses larger food subsidies (see the first row). Starting from a uniform tax policy with $\tau_a = 0$, a Rawlsian government is indifferent between optimizing food subsidies and increasing everyone's consumption aggregate by 0.07%.

The finding that food subsidies generate modest welfare gains does *not* mean that the welfare costs of setting suboptimal food subsidies are generally small. Figure 1 illustrates this point. It plots, for both a utilitarian and a Rawlsian criterion, the welfare costs of setting $\tau_a \in [-0.6, 0.6]$. To that end, we fix $\tau_a$ at a particular value on the horizontal axis and optimize with respect to the labor income tax $\tau_y$ and the lump-sum transfer $T$. The vertical axis then measures by how much welfare would increase if the government could also optimize

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

food subsidies. Naturally, the welfare costs are zero if $\tau_a$ is restricted at its optimal value, –4.68% for a utilitarian and –12.32% for a Rawlsian government. Furthermore, the costs of following a uniform consumption tax policy with $\tau_a = 0$ are 0.01% (0.07%) in consumption equivalent gains for a utilitarian (Rawlsian) government, see also Table 3. These numbers are substantially larger if food subsidies are severely mis-optimized. Specifically, a utilitarian (Rawlsian) government that faces the constraint $\tau_a = -0.6$ is indifferent between optimizing food subsidies (i.e., abandoning this constraint) and increasing everyone's consumption aggregate by 3.09% (2.42%).

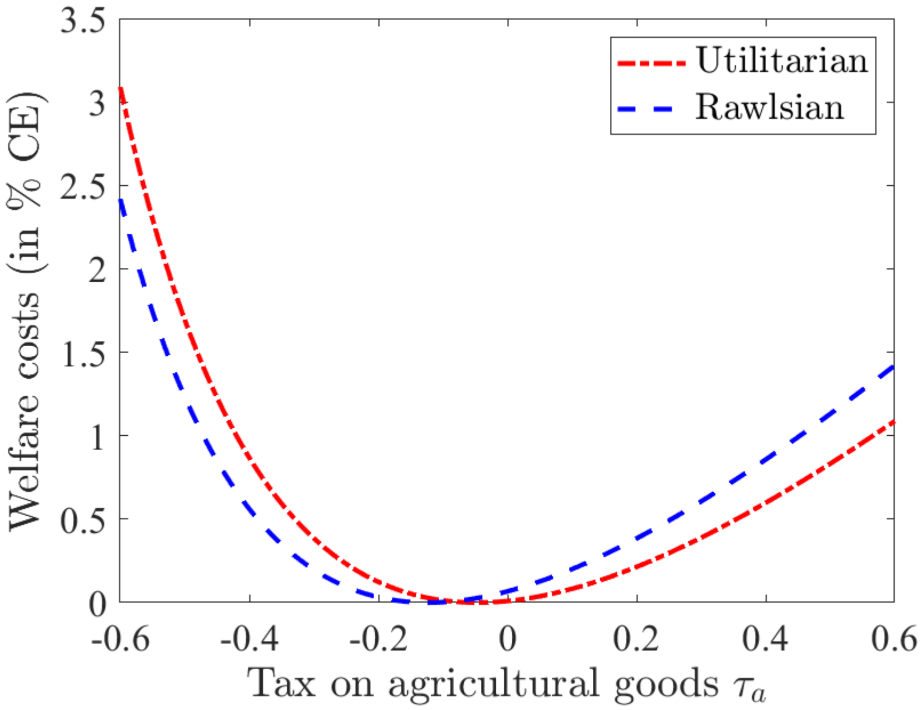



Figure 1: Welfare costs of suboptimal food subsidies

## 5.2 Robustness

Our optimal tax analysis indicates that general equilibrium effects rationalize food subsidies in the order of 5%–12% and that food subsidies generate modest welfare gains, in the range of increasing everyone's consumption aggregate by 0.01%–0.07%. This section investigates the robustness of these results by varying the elasticity of substitution i) between skill types in the production of both agricultural and non-agricultural goods ($\rho$) and ii) between agricultural goods and non-agricultural goods in the consumption aggregate ($\epsilon$).

#### 5.2.1 Strength of general equilibrium effects ($\rho$)

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 $\rho$. If it is very easy to substitute between skill types (i.e., if $\rho$ 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.

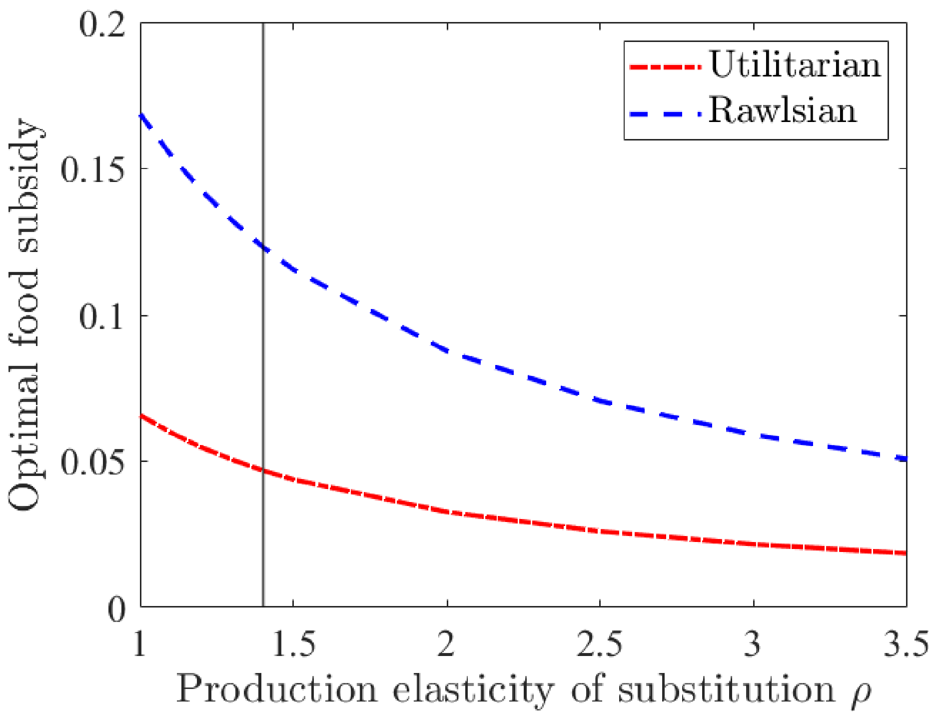



Figure 2: Optimal food subsidy for different values of $\rho$

Because our results are likely to be sensitive to this elasticity, we redo our analysis with different values of $\rho$. The remaining parameters are then set to match the moments outlined in Table 2. Figure 2 plots the optimal food subsidy for different values of $\rho \in [1.0, 3.5]$, both for a utilitarian and Rawlsian welfare function.[16] As can be seen from the figure, the

[16] The vertical line shows the baseline calibration with $\rho = 1.4$. In this case, the optimal food subsidy is 4.68% (12.32%) under the utilitarian (Rawlsian) criterion: see Table 3.

optimal food subsidy is declining in the elasticity of substitution $\rho$. This is intuitive: 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. Furthermore, for each value of $\rho$, the optimal food subsidy is larger under a Rawlsian than under a utilitarian objective. That is in line with the finding from Table 3. A Rawlsian planner cares only about the utility of the low-skilled workers. Consequently, it uses larger food subsidies in order to boost the wages of these workers. For plausible values of $\rho$ (from 1.0 to 2.5), the optimal food subsidy does not exceed 16.9% if the government has a Rawlsian objective and does not fall below 2.6% if the government has a utilitarian objective.

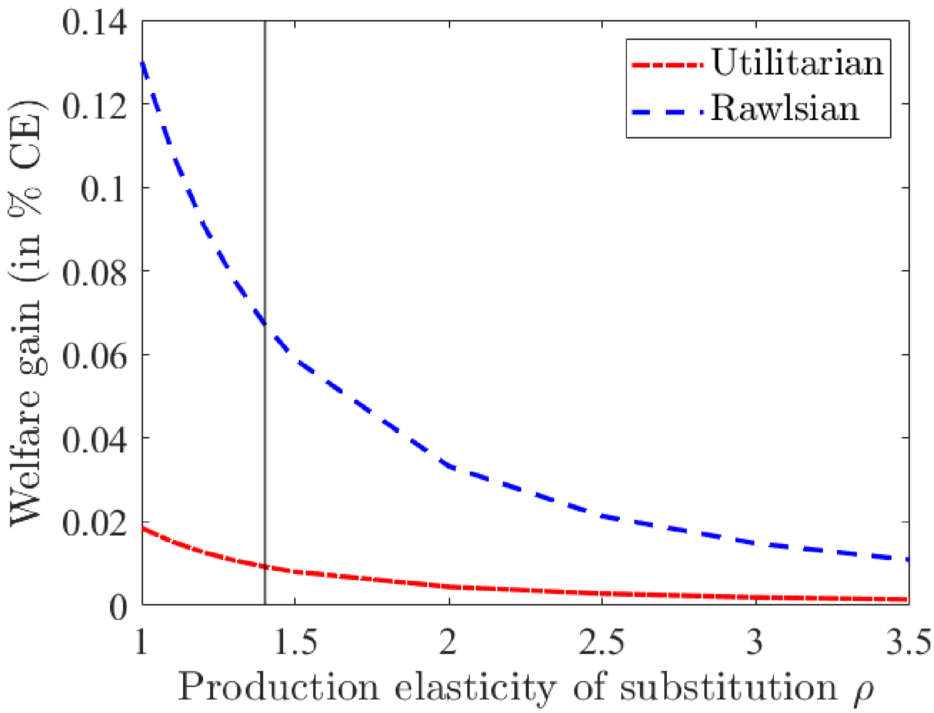



Figure 3: Welfare gains of food subsidies for different values of $\rho$

Figure 3 shows the welfare gains from using food subsidies for different values of $\rho$. It plots the percentage increase in the consumption aggregate that makes the government indifferent between this increase and optimizing food subsidies (starting from $\tau_a = 0$). The corresponding figures 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. Intuitively, there is less to gain from food subsidies if it is very easy to substitute between the different skill types. Moreover, the welfare gains are larger for a Rawlsian planner, who uses larger food subsidies in order to boost the wages of low-skilled workers. The welfare gains are modest in all cases: for all values of $\rho$ considered, optimizing food subsidies leads to a welfare gain equivalent to increasing everyone's consumption by at most 0.13%.

### 5.2.2 Strength of consumption responses ($\epsilon$)

The extent to which individuals change their consumption mix in response to a change in the food subsidy critically depends on $\epsilon$, 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 include 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 $\epsilon \in [0.3, 1.1]$. As before, the remaining parameters are set to match the moments from Table 2.

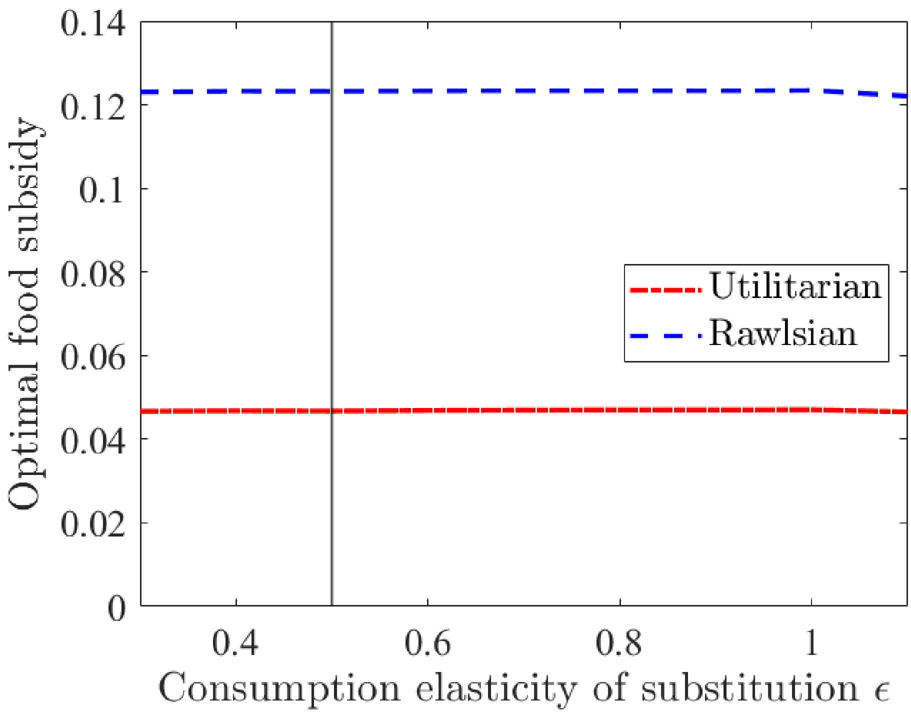



Figure 4: Optimal food subsidy for different values of $\epsilon$

Perhaps surprisingly, Figure 4 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 $\epsilon$ considered, the optimal food subsidy is slightly below 5% under a utilitarian criterion and slightly above 12% under a Rawlsian criterion – as in the baseline calibration with $\epsilon = 0.5$. The reason why the optimal food subsidy hardly responds to the degree of substitutability between agricultural and non-agricultural goods is that *both* the marginal costs and the marginal benefits of food subsidies are increasing in $\epsilon$. As explained earlier, a larger food subsidy leads to inefficiently high levels of food consumption, but also reduces the skill premium. By how much the skill premium declines depends critically on the increase in the demand for agricultural goods following a rise in the subsidy. The optimal food subsidy then strikes a balance between the marginal distributional benefits that come from general equilibrium effects and the marginal costs of distorting consumption decisions. Both the benefits and the costs 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 $\epsilon$.

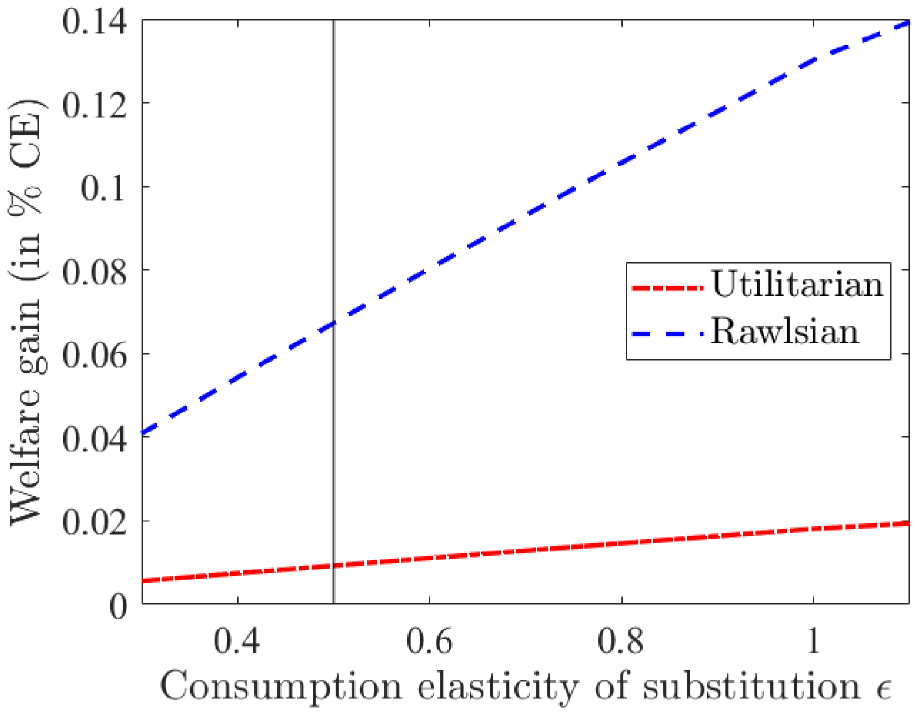



Figure 5: Welfare gains of food subsidies for different values of $\epsilon$

According to Figure 5, the welfare gains of using food subsidies are larger if it is easier to substitute between agricultural and non-agricultural goods. Recall that the welfare gains are calculated by comparing the unconstrained optimum to the constrained optimum (where $\tau_a = 0$). Welfare gains of food subsidies are increasing in $\epsilon$ for the following reason. When the food subsidy is optimal, the *marginal* distributional benefits that come from general equilibrium effects are equal to the *marginal* costs of distorting consumption decisions. 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. Naturally, 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. This is the case when it is easier to substitute between agricultural and non-agricultural goods. It follows that the welfare gains of using food subsidies increase in $\epsilon$. However, these welfare gains remain very modest throughout. For plausible values of $\epsilon$, they do not exceed 0.14% in consumption equivalents for the Rawlsian criterion, and are much smaller if the government has a utilitarian objective.

## 5.3 Nonlinear income taxes

So far, we have considered linear income tax schedules only: this is the most relevant case in countries with capacity constraints in taxation. We now ask what the optimal food subsidy is when the government can levy a *nonlinear* tax on labor income. This is the case theoretically analyzed in Naito (1999). Appendix C derives the formula for the optimal tax $\tau_a$ on agricultural products if the government levies a nonlinear income tax.

If we denote by $\mathcal{T}(\cdot)$ the tax schedule that is levied on labor income $w_i\ell_i$, the budget constraint of an individual with skill type $i \in \{L, H\}$ is

$$p_a(1+\tau_a)c_{a,i} + c_{n,i} = w_i\ell_i - \mathcal{T}(w_i\ell_i). \tag{22}$$

Clearly, the specification from equation (22) nests the baseline from Section 2 as a special case, with $\mathcal{T}(w_i\ell_i) = -T + \tau_y w_i\ell_i$. Solving the utility maximization problem leads to first-order conditions that are almost identical as before. The only difference is that $\tau_y$ in equation (5) is replaced by $\mathcal{T}'(w_i\ell_i)$, which potentially varies across skill types.

When the government can levy a nonlinear tax on labor income, it no longer faces the constraint that both skill types pay the same marginal tax rate. Instead, how much redistribution can be achieved is restricted by incentive constraints, see Mirrlees (1971) and Stiglitz (1982). The incentive constraints state that an individual of skill type $i \in \{L, H\}$ should not prefer to earn the same labor income (and hence, pay the same tax liability) as an individual of skill type $j \neq i$.[17] Formally, for both $i \neq j$, the following condition must hold:

$$u(c_{a,i}, c_{n,i}) + v(\ell_i) \geq u(c_{a,j}, c_{n,j}) + v(w_j\ell_j/w_i). \tag{23}$$

In words, the utility an individual of type $i$ attains when choosing her own bundle should be at least as large as the utility she attains when mimicking the other type. We verify numerically that only the incentive constraint of the high-skilled workers binds. Hence, what ultimately puts a constraint on how much the government can redistribute is that the high-skilled, who

[17]The fact that the utility function is separable between consumption and labor supply ensures that, if individuals with different skill types earn the same labor income, they also purchase the same amount of agricultural and non-agricultural goods.

earn a wage $w_H > w_L$, should not prefer to earn the same income as the low-skilled by reducing their labor supply.

We numerically solve the optimal tax problem and summarize the findings in Table 4. The parameterization is the same as in the baseline results (including a linear tax schedule). Focusing on the main results, we find that the optimal food subsidy is smaller than in the baseline calibration where the government levies a linear tax on labor income. It equals 1.42% if the government has a utilitarian objective and 5.94% if the government has a Rawlsian objective, compared to 4.68% and 12.32% with linear taxes. Apparently, the scope for using food subsidies to indirectly redistribute income through general equilibrium effects is smaller when the government can levy a nonlinear tax on labor income. In line with that finding, the welfare gains from using food subsidies are smaller as well. In particular, the welfare gain under a utilitarian criterion is a mere 1E-3% in consumption equivalents and 0.01% under a Rawlsian criterion, compared to 0.01% and 0.07% with linear taxes.

| | Utilitarian | Rawlsian |
|---|---|---|
| Tax on agricultural goods ($\tau_a$) | –1.42% | –5.94% |
| Marginal tax rate for the L-type ($\mathcal{T}'(w_L\ell_L)$) | 3.86% | 13.82% |
| Marginal tax rate for the H-type ($\mathcal{T}'(w_H\ell_H)$) | –6.58% | –27.23% |
| Consumption equivalent gain | 1E-3% | 0.01% |

Table 4: Optimal policy under a nonlinear income tax

The reason why food subsidies are less useful when the government can levy a nonlinear tax on labor income is that the government already uses income taxes to exploit general equilibrium effects for redistributive purposes. To see this, consider the marginal tax rates reported in the table. Under both a utilitarian and a Rawlsian criterion, the optimal marginal tax rate faced by low-skilled workers is positive, whereas the optimal marginal tax rate faced by high-skilled workers is negative.[18] This finding goes back to Stiglitz (1982), who formally shows these properties in a model with a single consumption good and endogenous wages that are derived from an aggregate production function.

To understand these properties of the optimal tax schedule, note that a positive marginal tax rate for low-skilled workers reduces their labor supply. This lowers their labor income, which makes it less attractive for a high-skilled worker to mimic a low-skilled worker. Furthermore, the reduction in labor supply also raises the wage $w_L$ of low-skilled workers. By contrast, the

[18]These properties do not say anything about the level of taxes, i.e., about average tax rates. For instance, if $G = \tau_a = 0$, low-skilled workers typically receive a transfer that is financed by taxes on high-skilled workers. In that case, the *average* tax rate faced by the low-skilled (high-skilled) is negative (positive), even though the *marginal* tax rate for this type is positive (negative).

negative marginal tax rate for high-skilled workers raises their labor supply, which lowers the wage $w_H$. Both forces contribute to a reduction in the skill premium. Specifically, at the optimal allocation with a nonlinear income tax, the skill premium is approximately 1.45 under a utilitarian criterion and 1.37 under the Rawlsian criterion (not displayed). A lower skill premium again makes it less attractive for a high-skilled worker to mimic a low-skilled worker by earning the same labor income. Hence, both the positive marginal tax rate for low-skilled workers and the negative marginal tax rate for high-skilled workers relax the incentive constraint (23), which raises welfare *ceteris paribus*.

The above discussion makes clear that, when the government has access to a nonlinear income tax on labor income, it already exploits general equilibrium effects for redistributive purposes. This significantly reduces the need for the government to use food subsidies to further reduce the skill premium. As a result, both the optimal food subsidy and the welfare gains from using them are much smaller when the government can levy a nonlinear tax on labor income. Despite this, we still find that there is a role for food subsidies, albeit smaller than in the baseline where the government can only levy a linear tax on labor income.

# 6 Tax reforms

The previous results concern the optimal tax system. We now connect these results to the three components identified in Proposition 1. Doing so highlights that general equilibrium effects are of the same magnitude as the more standard direct and behavioral effects. Thereafter, we use a similar decomposition to study the welfare effect of a reduction in the food subsidy, now starting from the *current* tax system. Our results suggest that ignoring general equilibrium effect can easily lead policymakers astray. Lastly, we ask whether it is possible to reform the current tax system in a Pareto-improving way. We find that it is, but that the scope for such improvements is limited.

## 6.1 Decomposing the optimal tax formula

The previous section demonstrated that the welfare gains from optimizing food subsidies are modest. This, however, does *not* mean that general equilibrium effects are small relative to the more standard direct and behavioral effects that often show up in optimal tax formulas, see for example Saez (2001) and Chetty (2009). To illustrate this, we consider a small increase in $\tau_a$ (i.e., a small reduction in the food subsidy), starting from the *optimal* tax system under either utilitarian or Rawlsian preferences. Recall from Proposition 1 that the welfare effect of an increase in $\tau_a$ can be decomposed as follows:

$$\frac{\partial \mathcal{W}}{\partial \tau_a}\frac{1}{\lambda} = DE + BE + GE. \tag{24}$$

Table 5 shows the impact on welfare that these direct, behavioral, and general equilibrium effects have – both for a utilitarian and a Rawlsian criterion. As the proposition requires, the sum of these effects is equal to zero at the optimal tax system. Unfortunately, there is no straightforward way to interpret these numbers.[19] One can, however, meaningfully interpret the sign of each component, as well as compare their magnitude.

| Criterion | $DE$ | $BE$ | $GE$ |
|---|---|---|---|
| Utilitarian | –0.05 | 0.26 | –0.21 |
| Rawlsian | –0.13 | 0.72 | –0.59 |

To facilitate readability, all effects are multiplied by a factor 100.

Table 5: Decomposing equation (18)

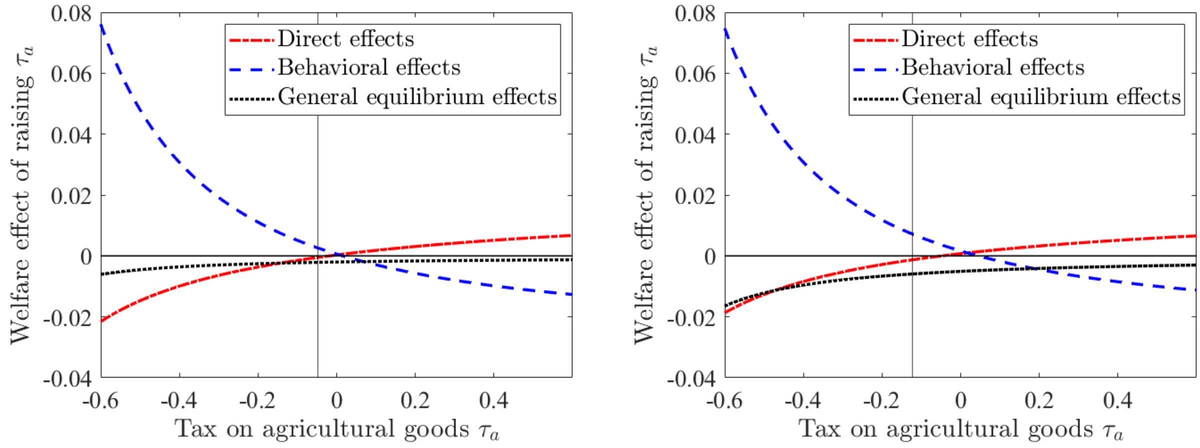
Under both specifications of the welfare function, the direct welfare effect of a reduction in the food subsidy is negative at the optimum. Hence, the distributional benefits of paying lower subsidies to high-skilled workers do not outweigh the costs of paying lower subsidies to low-skilled workers. The reason is that food enters the utility function as a necessity: $\underline{c}_a > 0$. A lower subsidy on a necessity disproportionately harms individuals with a lower income, i.e., the low-skilled. Considering the direct effect alone, lowering the subsidy on food thus comes at a welfare cost.

The welfare effect due to the behavioral responses is positive at the optimum, under both specifications of the welfare function. A lower subsidy on agricultural goods causes individuals to cut back on food consumption. Given that food is optimally subsidized, the reduction in food consumption by both low-skilled and high-skilled individuals generates a positive fiscal externality. The term $BE$ captures the associated welfare effect.

A reduction in the food subsidy also induces general equilibrium effects. Under both criteria, the associated impact on welfare is negative. This is intuitive: a higher $\tau_a$ (i.e., a lower food subsidy) reduces the demand for agricultural goods and thereby the demand for low-skilled workers. This, in turn, increases the skill premium. A higher skill premium has adverse distributional effects and thus leads to a reduction in welfare *ceteris paribus*. At the optimal tax system, the negative welfare impact of the general equilibrium effect exactly offsets the direct and behavioral effects.

The main insight from Table 5 is that the welfare effects due to general equilibrium responses

[19]The numbers are obtained from dividing a change in welfare $d\mathcal{W}$ (measured in utils) by a change in the tax rate $d\tau_a$, scaled by the multiplier $\lambda$ on the government budget constraint. To facilitate readability, all numbers are subsequently multiplied by a factor 100.




(a) Utilitarian criterion

(b) Rawlsian criterion

Figure 6: Decomposition of welfare gains

are of the same magnitude the direct and behavioral effects. Consequently, as we will demonstrate below, ignoring general equilibrium effects can easily lead policymakers astray when judging the desirability of a reduction in the food subsidy.

Figure 6 demonstrates how the direct, behavioral and general equilibrium effects jointly determine the optimal food subsidy. Starting at any value of $\tau_a \in [-0.6, 0.6]$, it plots the direct, behavioral and general equilibrium effects associated with a small reduction in the food subsidy. To do so, we first fix $\tau_a$ and optimize with respect the tax rate $\tau_y$ on labor income and the lump-sum transfer $T$. After that, we calculate the effects associated with a small increase in $\tau_a$. Naturally, at the optimal food subsidy the total welfare effect is equal to zero (this point is indicated by thin, vertical lines). By contrast, the total welfare effect of raising $\tau_a$ is positive (negative) at low (high) values of $\tau_a$.

Starting from uniform consumption taxes ($\tau_a = 0$), the general equilibrium effects make it worthwhile to marginally increase the subsidy (i.e., to lower $\tau_a$). As the food subsidy increases, however, the consumption distortions measured by the behavioral effects increase as well, at an increasing rate (Harberger, 1964). This makes further increases more and more costly. At the optimum, the behavioral effects have become large enough to cancel out the direct and general equilibrium effects, which change at a slower rate as $\tau_a$ is reduced. The main difference between the utilitarian and the Rawlsian criterion is that general equilibrium effects deliver larger welfare gains for a Rawlsian government, which makes for a larger optimal food subsidy.

## 6.2 Lowering the food subsidy

At the optimal tax system, the welfare impact of marginally increasing $\tau_a$ (i.e., reducing the food subsidy) is equal to zero. We now use the decomposition from equation (24) to study the welfare impact of lowering the food subsidy, starting from the *current* tax system.

To do so, we fix the lump-sum transfer $T$ at its current value. After that, we calculate the direct, behavioral and general equilibrium effects associated with a reduction in the food subsidy from its current level, both for a utilitarian and a Rawlsian criterion.[20] The reduction in the food subsidy is rebated through a lower income tax rate to make sure the government budget constraint is satisfied.[21] Because taxes are away from the optimum, the welfare effect of this reform is generally non-zero.

The results are shown in Table 6, for both the utilitarian and the Rawlsian criterion. As is the case at the optimal tax system (see Table 5), the direct and general equilibrium effects are negative, whereas the behavioral effects are positive. In line with what we found before, general equilibrium effects are of the same magnitude as the more standard direct and behavioral effects. Unlike before, however, the direct, behavioral and general equilibrium effects no longer sum to zero. Instead, a reduction in the current food subsidy that is rebated through a lower tax rate on labor income has a positive impact on welfare if the government has utilitarian preferences, but a negative impact if the government has Rawlsian preferences. This result should come as no surprise, as the current food subsidy lies in between these two.

| Criterion | $DE$ | $BE$ | $GE$ |
|---|---|---|---|
| Utilitarian | –0.17 | 0.57 | –0.19 |
| Rawlsian | –0.07 | 0.57 | –0.60 |

To facilitate readability, all effects are multiplied by a factor 100.

Table 6: Decomposing equation (24)

Table 6 also yields an example of how ignoring general equilibrium effects leads to wrong policy conclusions. We focus on the case of Rawlsian preferences. The sum of the direct and behavioral effects is positive: $DE + BE > 0$. Hence, ignoring general equilibrium effects

[20] To obtain the welfare weights, we solve the optimal tax problem under the additional constraints that both the food subsidy and the lump-sum transfer are kept at their current values.

[21] If instead of the lump-sum transfer $T$, the tax rate $\tau_y$ on labor income is kept fixed (see the previous footnote), the reduction in the food subsidy is rebated through a larger transfer. Because both the food subsidy and the lump-sum transfer at the optimal tax system under the utilitarian (Rawlsian) criterion are smaller (larger) than their current values, the results of that reform are less clear-cut.

one would conclude that a reduction in the food subsidy is welfare-enhancing. However, the total welfare effect of lowering the food subsidy is negative if general equilibrium effects are taken into account: $DE + BE + GE < 0$. Thus, a Rawlsian policymaker who calculates the first two components but ignores general equilibrium effects, will draw the wrong conclusion about which direction of tax reform is preferable.

## 6.3 Pareto-improving reforms

So far, we have calculated optimal taxes assuming the government has either a utilitarian or a Rawlsian objective. We finish by asking whether it is possible to reform the current tax system in a Pareto-improving way. To that end, we numerically solve for the tax system that maximizes the utility of a particular type (low-skilled or high-skilled), subject to the requirement that the utility of the other type is the same as in the calibrated economy.

Table 7 shows that it is possible to reform the tax system in a Pareto-improving way. Maximizing the utility of either type subject to keeping the other's utility fixed requires a small reduction in the food subsidy, a tiny reduction in the tax on labor income and a small increase in the transfer (compared to their current values of $\tau_a = -10\%$, $\tau_y = 14.53\%$ and $T/\bar{y} = 1.90\%$). The third column indicates that these reforms generate extremely small welfare gains. In particular, maximizing the utility of the low-skilled (high-skilled) subject to maintaining the other type's *current* utility level leads to a welfare gain of only 4E-5% (1E-3%) in consumption equivalents.

| | Improvement for L-type | Improvement for H-type |
|---|---|---|
| Tax on agricultural goods ($\tau_a$) | –9.72% | –9.71% |
| Tax on labor income ($\tau_y$) | 14.52% | 14.52% |
| Transfer/average income ($T/\bar{y}$) | 1.96% | 1.96% |
| Consumption equivalent gain | 4E-5% | 1E-3% |

Table 7: Pareto-improving policies

At this point, it should be emphasized that it is somewhat coincidental that the welfare gains are so small. The reason is that our calculations suggest that the current policy is not far from being Pareto optimal. To illustrate, suppose that $\alpha_L = 0.75$ and $\alpha_H = 1 - \alpha_L = 0.25$. Under these social preferences, the optimal tax rate on agricultural goods is $\tau_a = -9.75\%$ and the optimal tax rate on labor income is $\tau_y = 14.57\%$. These rates are Pareto optimal (since they maximize a Paretian objective), and very close to current policies.[22] We verify

[22] Of course, the finding that the current policy is close to one that maximizes a Paretian objective does *not* mean the current policy is close to maximizing social welfare. Making such a statement requires taking a stance on the Pareto weights $\alpha_L$ and $\alpha_H$, from which we refrain.

numerically that the scope for Pareto improvements is much larger if the current policy is farther away from one that maximizes a Paretian objective.[23]

# 7 Discussion and conclusion

This paper investigates to what extent food subsidies are helpful as a means to indirectly redistribute income. To do so, we analyze an economy with two types of labor (low-skilled and high-skilled) and two goods (agriculture and non-agriculture). In our model, 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 indirectly redistributes income from high-skilled to low-skilled workers. We derive a sufficient statistics formula for the optimal food subsidy that explicitly accounts for general equilibrium effects. Changes in equilibrium prices have both fiscal and distributional effects, which the government should take into account when designing food subsidies.

We then investigate the quantitative importance of this channel by calibrating our model to the economy of China, the world's largest subsidizer of agriculture (OECD, 2017). To make sure that any role for food subsidies is driven by general equilibrium effects, we assume preferences are separable between consumption and leisure and sub-utility of consumption is of the Stone-Geary form, with food entering as a necessity. 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.

Our findings suggest that general equilibrium effects rationalize food subsidies in the range 5%-12% depending on the redistributive preferences of the government, but that the welfare gains over uniform consumption taxation are small. We conduct further analysis and robustness exercises in a number of directions. On the whole, these confirm our main conclusions. Furthermore, we analyze a number of tax reforms. We show that general equilibrium effects are of the same magnitude as the more standard direct and behavioral effects that often show up in optimal tax formulas (Saez, 2001 and Chetty, 2009). Consequently, as we demonstrate, ignoring general equilibrium effects can lead policymakers astray when judging the desirability of tax reforms.

Our analysis abstracts from a number of important adjustment margins. We have assumed throughout that each worker's skill type is determined exogenously and the size of each skill group is fixed. Hence, a reduction in the skill premium driven by an increase in food subsidies

[23] The results are available upon request.

does not lead to different education choices. Allowing for an educational choice margin most likely reduces the impact of food subsidies on wages. This, in turn, limits the scope for the government to exploit general equilibrium effects for redistributive purposes (Saez, 2004). It is unclear how large the impact of such adjustments would be. What is clear, however, is that they are relevant only in the long run.

On another adjustment margin, we have taken the opposite stance. While each worker's skill type is fixed, she can freely move between sectors. One might also 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 a large-scale reallocation of employment between sectors (Herrendorf et al., 2014).

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 prices are determined on world markets. However, the price of agricultural goods in China appears very sensitive to country-specific shocks.[24] This is hardly surprising, considering that China is the world's largest producer of several agricultural goods. Furthermore, the vast majority of these goods are not exported but instead consumed domestically (FAO, 2012). For these reasons, we doubt that allowing for trade would significantly affect our results.[25]

Further research may clarify the role these adjustment margins play. In addition, it would be interesting to see how our quantitative results are affected if we allow for a more realistic income distribution. Adding these features contributes to understanding how general equilibrium effects shape optimal food subsidies, and consumption taxes more generally.

# References

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

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

[25]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.

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 (2020). Informality, consumption taxes and redistribution. NBER Working Paper No. 27429, 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.

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 (2020). Optimal redistribution with a shadow economy. mimeo, University of Bristol.

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

Feldstein, M. (1973). On the optimal progressivity of the income tax. *Journal of Public Economics 2*(4), 357–376.

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.

Harberger, A. (1964). The measurement of waste. *American Economic Review 54*(3), 58–76.

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.

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

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

Lopez, R., X. He, and E. De Falcis (2017). What drives China's new agricultural subsidies? *World Development 93*, 279–292.

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.

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.

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

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic*

*Studies 68*(1), 205–229.
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.
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 Optimal tax formulas

The government chooses tax policy $(\tau_a, \tau_y, T)$ to maximize welfare (12), subject to the budget constraint (13) and taking into account how equilibrium prices and quantities respond to changes in the tax instruments. In Appendix B we restate the conditions which pin down the equilibrium prices and quantities as a function of the tax instruments.

The Lagrangian of the government's optimization problem is

$$\begin{aligned}\mathcal{L} = &\sum_i \mu_i \alpha_i V_i(\tau_a, \tau_y, T) + \lambda \Bigg[\tau_a \sum_i \mu_i p_a(\tau_a, \tau_y, T) c_{a,i}(\tau_a, \tau_y, T) \\ &+ \tau_y \sum_i \mu_i w_i(\tau_a, \tau_y, T) \ell_i(\tau_a, \tau_y, T) - T - G\Bigg], \end{aligned} \tag{25}$$

where the summation takes place over skill types $i \in \{L, H\}$. We explicitly account for the dependency of the equilibrium prices and quantities on the tax instruments $(\tau_a, \tau_y, T)$. The indirect utility function $V_i$ is given by

$$\begin{aligned}V_i(\tau_a, \tau_y, T) = \max_{c_{a,i}, c_{n,i}, \ell_i, \eta_i} &\Bigg\{ u(c_{a,i}, c_{n,i}) + v(\ell_i) \\ &+ \eta_i \Big[T + w_i(\tau_a, \tau_y, T)(1 - \tau_y)\ell_i - c_{n,i} - p_a(\tau_a, \tau_y, T)(1 + \tau_a) c_{a,i}\Big]\Bigg\}, \end{aligned} \tag{26}$$

where $\eta_i$ is the multiplier on the individual budget constraint. From the first-order condition with respect to $c_{n,i}$, at the optimum the latter coincides with the marginal utility of non-agricultural consumption: $u_{n,i} = \eta_i$. To determine how a change in the tax instruments affects individual utility $V_i$, differentiate equation (26) with respect to the policy instruments. By the envelope theorem,

$$\frac{\partial V_i}{\partial \tau_a} = \eta_i \Big[ - p_a c_{a,i} + (1 - \tau_y)\ell_i \frac{\partial w_i}{\partial \tau_a} - (1 + \tau_a) c_{a,i} \frac{\partial p_a}{\partial \tau_a}\Big], \tag{27}$$

$$
\frac{\partial V_i}{\partial \tau_y} = \eta_i \bigg[ - w_i \ell_i + (1-\tau_y)\ell_i \frac{\partial w_i}{\partial \tau_y} - (1+\tau_a) c_{a,i} \frac{\partial p_a}{\partial \tau_y} \bigg], \tag{28}
$$

$$
\frac{\partial V_i}{\partial \tau_a} = \eta_i \bigg[ 1 + (1-\tau_y)\ell_i \frac{\partial w_i}{\partial T} - (1+\tau_a) c_{a,i} \frac{\partial p_a}{\partial T} \bigg]. \tag{29}
$$

Turning to the optimal tax problem, the first-order conditions associated with the Lagrangian (25) are:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \tau_a} =& \sum_i \mu_i \alpha_i \frac{\partial V_i}{\partial \tau_a} + \lambda \bigg[ \sum_i \mu_i p_a c_{a,i} + \bigg( \tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial \tau_a} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial \tau_a} \bigg) \\
&+ \bigg( \tau_a \sum_i \mu_i c_{a,i} \frac{\partial p_a}{\partial \tau_a} + \tau_y \sum_i \mu_i \ell_i \frac{\partial w_i}{\partial \tau_a} \bigg) \bigg] = 0, \qquad (30) \\
\frac{\partial \mathcal{L}}{\partial \tau_y} =& \sum_i \mu_i \alpha_i \frac{\partial V_i}{\partial \tau_y} + \lambda \bigg[ \sum_i \mu_i w_i \ell_i + \bigg( \tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial \tau_y} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial \tau_y} \bigg) \\
&+ \bigg( \tau_a \sum_i \mu_i c_{a,i} \frac{\partial p_a}{\partial \tau_y} + \tau_y \sum_i \mu_i \ell_i \frac{\partial w_i}{\partial \tau_y} \bigg) \bigg] = 0, \qquad (31) \\
\frac{\partial \mathcal{L}}{\partial T} =& \sum_i \mu_i \alpha_i \frac{\partial V_i}{\partial T} + \lambda \bigg[ -1 + \bigg( \tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial T} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial T} \bigg) \\
&+ \bigg( \tau_a \sum_i \mu_i c_{a,i} \frac{\partial p_a}{\partial T} + \tau_y \sum_i \mu_i \ell_i \frac{\partial w_i}{\partial T} \bigg) \bigg] = 0, \qquad (32) \\
\frac{\partial \mathcal{L}}{\partial \lambda} =& \tau_a \sum_i \mu_i p_a c_{a,i} + \tau_y \sum_i \mu_i w_i \ell_i - T - G = 0. \qquad (33)
\end{aligned}
$$

Combined, these equations pin down optimal tax policy $(\tau_a, \tau_y, T)$ together with the multiplier $\lambda$ on the government budget constraint.

To derive the optimal tax formulas, substitute out for the impact of the tax instruments on individual utility using equations (27)–(29) in equations (30)–(32) and use the first-order condition from the individual maximization problem $\eta_i = u_{n,i}$. Rearranging gives:

$$
\begin{aligned}
&\sum_i \mu_i (\lambda - \alpha_i u_{n,i}) p_a c_{a,i} + \lambda \bigg( \tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial \tau_a} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial \tau_a} \bigg) \\
&+ \bigg( \sum_i \mu_i (\lambda \tau_a - \alpha_i u_{n,i}(1+\tau_a)) c_{a,i} \frac{\partial p_a}{\partial \tau_a} + \sum_i \mu_i (\lambda \tau_y + \alpha_i u_{n,i}(1-\tau_y)) \ell_i \frac{\partial w_i}{\partial \tau_a} \bigg) = 0, \qquad (34) \\
&\sum_i \mu_i (\lambda - \alpha_i u_{n,i}) w_i \ell_i + \lambda \bigg( \tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial \tau_y} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial \tau_y} \bigg) \\
&+ \bigg( \sum_i \mu_i (\lambda \tau_a - \alpha_i u_{n,i}(1+\tau_a)) c_{a,i} \frac{\partial p_a}{\partial \tau_y} + \sum_i \mu_i (\lambda \tau_y + \alpha_i u_{n,i}(1-\tau_y)) \ell_i \frac{\partial w_i}{\partial \tau_y} \bigg) = 0, \qquad (35) \\
&\sum_i \mu_i (\alpha_i u_{n,i} - \lambda) + \lambda \bigg( \tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial T} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial T} \bigg) \\
&+ \bigg( \sum_i \mu_i (\lambda \tau_a - \alpha_i u_{n,i}(1+\tau_a)) c_{a,i} \frac{\partial p_a}{\partial T} + \sum_i \mu_i (\lambda \tau_y + \alpha_i u_{n,i}(1-\tau_y)) \ell_i \frac{\partial w_i}{\partial T} \bigg) = 0. \qquad (36)
\end{aligned}
$$

To arrive at equation (18) from Proposition 1, divide equation (34) by $\lambda$ and use the definition of the welfare weights $g_i = \frac{\alpha_i u_{n,i}}{\lambda}$:

$$\underbrace{\sum_i \mu_i(1-g_i)p_a c_{a,i}}_{DE} + \underbrace{\left(\tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial \tau_a} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial \tau_a}\right)}_{BE}$$
$$+ \underbrace{\left(\sum_i \mu_i(\tau_a - g_i(1+\tau_a))c_{a,i}\frac{\partial p_a}{\partial \tau_a} + \sum_i \mu_i(\tau_y + g_i(1-\tau_y))\ell_i \frac{\partial w_i}{\partial \tau_a}\right)}_{GE} = 0. \tag{37}$$

A similar expression can be derived for the optimal tax on labor income $\tau_y$ and the optimal transfer $T$ by dividing equations (35)–(36) by $\lambda$:

$$\sum_i \mu_i(1-g_i)w_i\ell_i + \left(\tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial \tau_y} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial \tau_y}\right)$$
$$+ \left(\sum_i \mu_i(\tau_a - g_i(1+\tau_a))c_{a,i}\frac{\partial p_a}{\partial \tau_y} + \sum_i \mu_i(\tau_y + g_i(1-\tau_y))\ell_i \frac{\partial w_i}{\partial \tau_y}\right) = 0, \tag{38}$$
$$\sum_i \mu_i(g_i - 1) + \left(\tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial T} + \tau_y \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial T}\right)$$
$$+ \left(\sum_i \mu_i(\tau_a - g_i(1+\tau_a))c_{a,i}\frac{\partial p_a}{\partial T} + \sum_i \mu_i(\tau_y + g_i(1-\tau_y))\ell_i \frac{\partial w_i}{\partial T}\right) = 0. \tag{39}$$

As stated earlier, together with the government budget constraint (33), the optimal tax formulas (37)–(39) jointly pin down the optimal policy $(\tau_a, \tau_y, T)$ and the multiplier $\lambda$. These equations feature the impacts of the tax instruments on equilibrium prices and quantities. Below we restate the conditions that can be used to determine these effects.

# B Equilibrium given tax policy

The following conditions pin down the equilibrium consumption and labor supply decisions $\{(c_{a,L}, c_{n,L}, \ell_L), (c_{a,H}, c_{n,H}, \ell_H)\}$, labor inputs $\{(L_a, H_a), (L_n, H_n)\}$ and prices $(p_a, w_L, w_H)$ for a given tax policy $(\tau_a, \tau_y, T)$. Hence, these equations can be used to determine how equilibrium prices and quantities vary with the tax instruments. These are the partial effects that show up in the optimal tax formulas (37)–(39).

From the household maximization problem, for each $i \in \{L, H\}$,

$$\frac{C(c_{a,i}, c_{n,i})^{-\sigma}}{P} = \frac{\psi(1-\ell_i)^{-\phi}}{w_i(1-\tau_y)}, \tag{40}$$
$$c_{n,i} = \frac{1-\omega}{\omega}(c_{a,i} - \underline{c}_a)(p_a(1+\tau_a))^{\epsilon}, \tag{41}$$
$$p_a(1+\tau_a)c_{a,i} + c_{n,i} = T + w_i(1-\tau_y)\ell_i, \tag{42}$$

where $C(c_{a,i}, c_{n,i})$ 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

$$L_a + L_n = \mu_L \ell_L, \tag{45}$$

$$H_a + H_n = \mu_H \ell_H, \tag{46}$$

$$Y_a = \mu_L c_{L,a} + \mu_H c_{H,a}. \tag{47}$$

Combined, equations (40)–(47) constitute a system of 13 equations in 13 unknowns (recall: there are two sectors and skill types), which pin down the equilibrium quantities and prices for a given tax policy $(\tau_a, \tau_y, T)$. As such, this system can be used to determine the impact of the tax instruments on equilibrium outcomes which show up in the optimal tax formulas (37)–(39), for example by implicit differentiation. Doing so, however, leads to complicated expressions involving many terms. For this reason, we prefer to express the optimal tax formulas in terms of direct, behavioral and general equilibrium effects.

It is worth pointing out that $G$ does not show up in the above system. Hence, for a particular choice of the tax instruments $(\tau_a, \tau_y, T)$, 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 as well. This is why in the optimal tax problem studied above, 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 quantities on tax policy $(\tau_a, \tau_y, T)$.

# C Optimal food subsidy with a nonlinear income tax

Suppose the government can levy a nonlinear tax $\mathcal{T}(\cdot)$ on labor income $w_i \ell_i$, which may include a lump-sum transfer. The equilibrium given tax policy is then described by equations (40)–(47), with equations (40) and (42) replaced by, respectively,

$$\frac{C(c_{a,i}, c_{n,i})^{-\sigma}}{P} = \frac{\psi(1-\ell_i)^{-\phi}}{w_i(1-\mathcal{T}'(w_i \ell_i))}, \tag{48}$$

$$p_a(1+\tau_a)c_{a,i} + c_{n,i} = w_i \ell_i - \mathcal{T}(w_i \ell_i). \tag{49}$$

See also the discussion in Section 5.3. Together, equations (41) and (43)–(49) pin down all equilibrium quantities and prices given, among other things, the tax $\tau_a$ on agricultural goods. The government budget constraint (13), in turn, is replaced by

$$\tau_a \sum_i \mu_i p_a c_{a,i} + \sum_i \mu_i \mathcal{T}(w_i \ell_i) = G. \tag{50}$$

The Lagrangian of the government's optimization problem then reads

$$\mathcal{L} = \sum_i \mu_i \alpha_i V_i(\tau_a) + \lambda \left[ \tau_a \sum_i \mu_i p_a(\tau_a) c_{a,i}(\tau_a) + \sum_i \mu_i \mathcal{T}(w_i(\tau_a)\ell_i(\tau_a)) - G \right], \tag{51}$$

where we highlight the dependency of equilibrium prices and quantities on $\tau_a$. Moreover, the indirect utility function is

$$\begin{aligned} V_i(\tau_a) = \max_{c_{a,i}, c_{n,i}, \ell_i, \eta_i} \Big\{ & u(c_{a,i}, c_{n,i}) + v(\ell_i) \\ & + \eta_i \Big[ w_i(\tau_a)\ell_i - \mathcal{T}(w_i(\tau_a)\ell_i) - c_{n,i} - p_a(\tau_a)(1+\tau_a)c_{a,i} \Big] \Big\}, \end{aligned} \tag{52}$$

which is the counterpart of equation (26). The formula for the optimal tax $\tau_a$ on agricultural goods can then be derived using exactly the same steps as in Appendix A. The corresponding optimal tax formula is

$$\begin{aligned} & \underbrace{\sum_i \mu_i (1-g_i) p_a c_{a,i}}_{DE} + \underbrace{\left( \tau_a \sum_i \mu_i p_a \frac{\partial c_{a,i}}{\partial \tau_a} + \sum_i \mu_i w_i \frac{\partial \ell_i}{\partial \tau_a} \mathcal{T}'(w_i \ell_i) \right)}_{BE} \\ & + \underbrace{\left( \sum_i \mu_i (\tau_a - g_i(1+\tau_a)) c_{a,i} \frac{\partial p_a}{\partial \tau_a} + \sum_i \mu_i (\mathcal{T}'(w_i\ell_i) + g_i(1-\mathcal{T}'(w_i\ell_i)))\ell_i \frac{\partial w_i}{\partial \tau_a} \right)}_{GE} = 0. \end{aligned} \tag{53}$$

The only difference between this result and the optimal tax formula (37) when the government levies a linear tax on labor income (which is also stated in Proposition 1) is that the marginal tax rate on labor income $\tau_y$ is replaced by $\mathcal{T}'(w_i \ell_i)$, which potentially varies across skill types. Intuitively, both the fiscal externalities associated with wage and labor supply responses as well as the impact of a change in the wage on indirect utility depend on the marginal tax rate $\mathcal{T}'(w_i \ell_i)$.