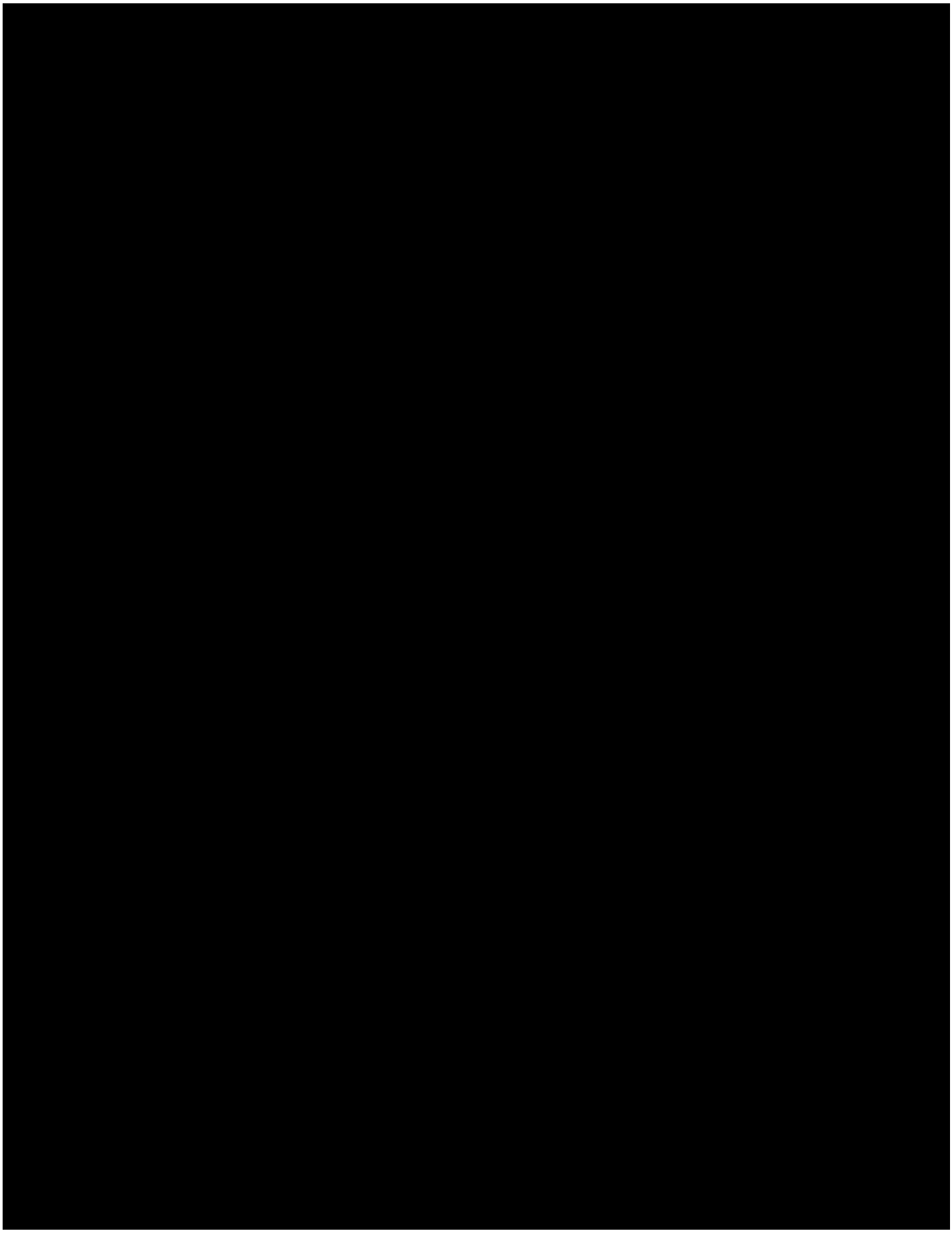

 WP/18/146

IMF Working Paper

Strategy, Policy, and Review Department

The Welfare Implications of Fiscal Consolidations in Low-income Countries

Prepared by Adrian Peralta-Alva, Xuan S. Tam, Xin Tang and Marina M. Tavares*

Authorized for distribution by Johannes Wiegand

June 2019

Abstract

We quantitatively investigate the welfare costs of fiscal consolidations in low-income countries through value added tax (VAT), personal income tax (PIT), and corporate income tax (CIT). We extend the standard heterogeneous agents incomplete markets model by including multiple sectors and regions to capture salient features of low-income countries. We find that VAT has the least efficiency costs but reduces welfare substantially by widening the rural-urban gap. Despite causing larger output loss, PIT and CIT have less welfare costs because the tax incidence is distributed evenly between regions. For all the taxes considered, uninsurable idiosyncratic risks cause sizable distributional costs.

JEL Classification Numbers: D31, E62, H23

Keywords: Fiscal Consolidations, Welfare Decomposition, Low-income Countries

Author's E-mail Addresses: aperalta-alva@imf.org, mmendestavares@imf.org, xuanstam@cityu.edu.hk, xtang@imf.org.

* Peralta-Alva, Tang, Tavares: International Monetary Fund; Tam: City University of Hong Kong. This paper is a substantially revised version of the previous version titled "The Macroeconomic and Distributional Implications of Fiscal Consolidations in Low-income Countries." For helpful comments we thank Alexis Anagnostopoulos, Marina Azzimonti, Andrew Berg, Yongsung Chang, Juan Carlos Conesa, Rupa Duttagupta, Ashique Habib, Hui He, Jonathan Heathcote, Gabriel Mihalache, Catherine Pattillo, Chris Papageogiou, Rishabh Sinha, David Wiczer, Xican Xi and participants various seminars. This paper is part of a research project on macroeconomic policy in low-income countries supported by the Department for International Development (DFID) of the United Kingdom. The authors thank the DFID for generous financial support. The research results and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the DFID of the United Kingdom.

TABLE OF CONTENTS

FIGURES

TABLES

I. Introduction

Low-income countries have low tax revenue to GDP ratio. The average tax to GDP ratio in low-income countries is 15% compared to that of 30% in advanced economies. Meanwhile, these countries are also those in most need of fiscal space for sustainable and inclusive growth. In the past two decades, low-income countries have made substantial efforts in strengthening revenue mobilization. There is a grand literature providing guidance to tax reforms in advanced economies, but the literature with a special focus on low-income countries is thin. The purpose of this paper is thus to study quantitatively the welfare costs of fiscal consolidations using different tax instruments in the context of low-income countries.

The economic structure of low-income countries differs substantially from that of advanced economies. In particular, low-income countries usually have a large unproductive agricultural sector, face serious informality issue, lack instruments to insure against idiosyncratic income risks, and observe a sharp rural-urban distinction.[1] These features alter how the economy respond to different taxes. Therefore, to guide our analysis, we extend the workhorse heterogeneous agents incomplete markets model [Aiyagari (1994)] to include necessary elements to cope appropriately with the economic structure of low-income countries.

Specifically, we consider a model with four sectors of different productivity levels—food, manufacturing, services, exported cash crops—two regions with segmented labor markets—rural and urban—and a unified capital market. Each region is populated by a continuum of households who consume food, manufacturing goods, and services. Each household also faces persistent idiosyncratic productivity shocks that can only be partially insured against using one period risk-free bond. Based on their comparative advantage, households divide their total hours between the formal and informal markets in their dwelling region [Roy (1951)]. The formal and informal labor markets in each region host different sectors. The agricultural sectors are hosted exclusively in the rural area, where workers in the formal and informal markets are hired to produce respectively cash crops and food. On the other hand, manufacturing goods and services are provided respectively by urban households who work in the formal and informal markets. A utilitarian government has access to three Ramsey taxes: value added tax (VAT) on food and manufacturing goods consumption, personal income tax (PIT) on formal wage income, and corporate income tax (CIT) on revenues of manufacturing firms and profits from cash crops production.

To discipline our quantitative analysis, we calibrate our benchmark economy to Ethiopia, which we choose as representative of low-income countries.[2] Our model replicates well the sectoral consumption and production shares, tax structure, and regional consumption inequality observed in

[1] These features of low-income countries are well documented in the literature. See for instance, Townsend (1994), Udry (1994), Schneider and Enste (2000), Lagakos and Waugh (2013), Young (2013), Gollin, Lagakos and Waugh (2014), Herrendorf, Rogerson and Valentinyi (2014), and La Porta and Shleifer (2014) on these various aspects of low-income countries.

[2] Thanks to several IMF country teams that dedicated to using the model for policy advices, we were able to verify that the main messages from the paper carry through a number of developing economies. Specifically, the model has been calibrated and applied to the case of Benin, Cambodia, Dominican Republic, Ethiopia, Philippines, Senegal, and Serbia. The corresponding Article IV consultations reports can be downloaded from the official website of the International Monetary Fund.

Ethiopia. We then use the calibrated model to evaluate the welfare costs of raising additional tax revenues of 2% GDP in the benchmark economy using VAT, CIT and PIT respectively. The tax reforms are assumed to be unexpected and permanent.

Our quantitative results show that across the three taxes, VAT has the least efficiency costs. Steady state comparison shows that VAT leads to a 1.5% loss in output, while CIT and PIT cause respectively 4.9% and 6.2% decline. However, the consumption equivalence of the total welfare costs of VAT (3.89%) is the highest among all the reforms, with PIT (3.31%) and CIT (2.24%) follow in order. Decomposition of the welfare costs shows that the distributional component (1.32%) is what makes VAT particularly undesirable comparing to PIT and CIT, for which the components are small. Closer inspection reveals that this is mainly because VAT substantially worsens the urban-rural gap, in that the welfare cost born by the rural households is 8 times that by urban households. The reason is that in our model, the government expenditure is spent on manufacturing goods produced in the urban area. While the tax incidence falls on households in both area when they consume, some of the tax revenues are implicitly rebated to urban households through government demand on manufacturing goods, reducing its overall impact on urban households. In a nutshell, VAT widens the urban-rural gap because it has the side effect of transferring income from rural to urban households. In addition, the aggregate component of VAT (2.61%) is higher than CIT (2.52%), though CIT causes nearly twice as much drop in aggregate consumption. The discrepancy is caused by first order distortion on after-tax prices and hence the optimal bundle of the consumption goods since VAT is not imposed on informal services.

Overall, we find that the major costs of VAT is the cross-region distribution of tax incidence, while those of CIT and PIT are losses in aggregate consumption. This suggests that VAT is best accompanied by transfer programs, and pro-growth policy works better with CIT and PIT. The aforementioned results suggest that both the aggregate and distributional components of tax reforms in low-income countries differ from those in advanced economies nontrivially, and hence policy prescriptions derived from standard models are in the risk of being misplaced.

Pioneered by Domeij and Heathcote (2004), most papers in the literature found that the substantial redistribution during transition makes steady-state welfare comparisons misleading.[3] Contrary to what has been found in the literature, we find that for low-income countries, transitional dynamics are less important because low household savings (or equivalently low capital stock) lead to fast convergence to the new equilibrium. For the same reason, on the regional level, short-run transition matters more to the urban households, as they hold more savings relative to the rural households. Quantitatively, the welfare costs of VAT are larger when transitional dynamics are considered, while those of CIT and PIT are smaller. This result is because both CIT and PIT reduce the steady state capital stock, allowing households to enjoy more consumption during the transition.

Studies in development economics have firmly established that households in developing countries face severe difficulty to insulate themselves from volatility in income [Townsend (1994) and Udry (1994)].[4] For this reason, we further explore how idiosyncratic risks affect the welfare costs of the tax reforms. We find that eliminating idiosyncratic risks results in substantial expansion of

[3] For recent work, see for example Anagnostopoulos, Cárceles-Poveda and Lin (2012), Bakış, Kaymak and Poschke (2015), and Conesa, Li and Li (2018).

[4] See Donovan (2018) and Morten (2019) for two recent examples.

the size of the economy. Specifically, output and capital stock are doubled, and labor supply to the formal sector is almost tripled. This finding contrasts those in the precautionary saving literature. The reason is that in absence of risks, households work more in the formal sector to take advantage of the higher productivity, while they hesitate to do so before to control the overall risk in their income. Short-run dynamics here are more important because first the transition takes longer with higher capital stock, and second the interest rate and urban wage are constant across steady states but are not during transitions. Comparing the welfare costs of tax reforms across the economies with and without idiosyncratic risks, we find that the aggregate components are similar for both scenarios, but the distributional components differ substantially.[5] This happens because in the no risk case, households are wealthier and consequently, saving takes up a higher fraction in household's total disposable income. As a result, reduction in the prices of consumption goods equalizes the indirect utility by more given the concavity of the utility function. While in the benchmark case, the drops in consumption goods prices would also cut the income of poor households with little saving, leaving consumption largely intact.

Related Literature.—This paper is most closely related to the extensive literature studying taxation quantitatively in heterogeneous agents incomplete markets models. Important work in this literature includes Aiyagari (1995), Domeij and Heathcote (2004), Conesa, Kitao and Krueger (2009), and so forth. These papers study the distribution of tax incidence by changing the mix of consumption, labor and capital income tax for the United States. Our paper shares a similar theme, but extends the discussion to the environment of low-income countries, which constitutes areas that differ in terms of industry and income level. This allows us to quantify how tax incidence is distributed across regions which cannot be studied with standard one sector model. Our finding that VAT implicitly redistributes income to the urban households shows that this additional granularity indeed matters. In fact, although previous studies have also highlighted the regressiveness of consumption tax as its major drawback, we show that for low-income countries it operates through an additional channel of cross-region redistribution which is novel to the literature. A flip-side of the result is that CIT and PIT can be much less costly to use due to a more uniform distribution of tax incidence across regions.

Moreover, most studies find that the redistribution along the transition path has large welfare effects that make steady-state comparisons misleading. We show that since low-income countries are mostly low on capital stock, the transition to the new equilibrium with a local change in tax rate is usually fast, allowing calculations from steady-state comparisons to be reasonable approximations.

The way idiosyncratic risks interact with the multi-region multi-sector structure of our model connects our paper to several other strands of literature. The lack of precautionary saving when income from the formal sector is more volatile contributes to the discussion of labor supply under incomplete markets [Flodén (2006), Pijoan-Mas (2006), and Marcet, Obiols-Homs and Weil (2007)].[6] We show that idiosyncratic risks affect household's labor supply decision and conse-

[5] For the same reason as in Domeij and Heathcote (2004), we scale up the equilibrium wealth distribution in the benchmark case such that the aggregate capital equals to the level in the case of no idiosyncratic risks.

[6] In a loose sense, the formal-informal occupational choice decision in our paper is also related to an earlier literature in macroeconomics and development economics that emphasize the importance of home production. See for instance, Benhabib, Rogerson and Wright (1991), Parente, Rogerson and Wright (2000), and Gollin, Parente and Rogerson (2004).

quently capital stock and sectoral output in a way similar to the risk-return management in a portfolio choice model [Angeletos (2007), Allen and Atkin (2017), and Donovan (2018)]. Our findings suggest that by increasing formal labor supply and capital accumulation, reducing or providing better insurance to labor market shocks in the context of low-income countries could potentially bring larger welfare gains than those found in advanced economies.[7] Also, we would like to emphasize that not only does reducing labor market shocks itself brings (expected) welfare gains, it also makes revenue mobilizations less costly. Moreover, the fact that in of our model, the efficiency costs of CIT is lower than those of VAT relates to the optimal taxation literature by showing that distortion to the optimal consumption bundle could provide new insights to the classical Chamley-Judd results.[8]

The remainder of the paper proceeds as follows. We first describe the model in Section I and calibrate its benchmark version in Section II. Section III contains the main results, and we conclude in Section IV. All technical details are deferred to the appendices.

II. The Model

We extend the standard Aiyagari (1994) heterogenous agents incomplete markets model to capture salient features of low-income economies, namely a large agricultural sector, significant informal economy, and a distinction between urban and rural areas.

Take Ethiopia as an example of a typical developing country. Ethiopia has a purchasing power parity adjusted GDP per capita of about 3% of the U.S level, making it one of the poorest countries in the world. It has a large and unproductive agricultural sector, which employs about 70% of the work force. Agricultural production in Ethiopia is overwhelmingly of a subsistence nature, and a large part of commodity exports are provided by the small agricultural cash crops sector, with coffee as the largest foreign exchange earner, and its flower industry becoming a new source of revenue in recent years. Our model attempts to capture these features in a stylized way.

A. *The Environment*

Our model is a small open economy with two regions—rural and urban—and four sectors—domestic and exporting agriculture (later referred to as food and cash crops), manufacturing, and services.[9, 10] Each region is populated by a continuum of infinitely lived households. The popula-

[7] See Heathcote, Storesletten and Violante (2008, 2014, 2017) and Kaplan and Violante (2010) for several recent contributions on quantifying consumption insurance and its welfare implications.

[8] For established work in this literature, see Judd (1985), Chamley (1986), Coleman (2000), and Werning and Straub (2018). Chari and Kehoe (1999) provide a review of early contributions.

[9] As will be explained in the calibration section, these sectors should not be interpreted by their literal names. By manufacturing and services, what we really mean is goods and services that produced formally or informally in the urban area. For instance, airline, telecommunication, and modern financial services are mapped to the manufacturing as opposed to services sector in our model. Hence what differentiates them is not their statistical labels, but the way production is organized, here in particular, the production functions.

[10] To clarify, the exogenous price in our small open economy is that of cash crops. Usually in the international economics literature, interest rate is the exogenous price [for instance, Mendoza, Quadrini and Ríos-Rull (2009)]. In our model, interest rate is endogenously determined domestically.

tion shares for rural and urban areas are μ^r and μ^u. Each household is endowed with one unit of time, which is divided between the regional formal labor market and informal self-employment. We assume that labor markets in the two regions are segmented, and households cannot migrate.[11]

The agricultural sectors are exclusively hosted in the rural area, with the manufacturing and services sectors housed in the urban area. More specifically, in the rural area, households work informally on their own arable land to grow food. A share μ^f of large farmers hire labor in the formal market to produce both food and cash crops. We use the large farmer to model in a parsimonious way the large-scale agricultural exporting sector and seasonal hiring during the labor intensive stages of agricultural production (e.g., planting and harvesting) which rely on temporarily hired labor.[12] In the urban area, households work informally to provide services and a representative neoclassical firm hires labor in the formal market together with rented capital to produce manufacturing goods. We assume that urban households face idiosyncratic productivity shocks in the manufacturing sector, while the shocks hit rural households when they work in their own plot.[13] With the combined formal and informal income, households make a consumption-saving decision where they have access to one risk-free asset and divide their total consumption expenditure over food, manufacturing goods, and services optimally. We maintain the standard assumption in the literature that savings are turned into capital of equivalent value. Further, in our model, cash crops production is modernized in the sense that it also employs capital, meaning that capital is used in both manufacturing goods and cash crops production. Both capital depreciates at rate δ. Large farmers' income sources are revenues from selling domestic and exporting agricultural goods. The income is used to finance consumption and investment in machinery. Large farmers possesses their own capital. We let the manufacturing goods be the numeraire, and p^a and p^s be the relative price of food and services. The wage rates are w^m and w^f respectively for urban and rural formal market. The risk-free asset yields a return of r.

We assume that food and services are used exclusively for domestic consumption, and cash crops serve only the international market. We assume that in each period, the current account is balanced by the government through importing manufacturing goods.[14] The manufacturing goods are used for consumption and capital. This broadly captures the pattern that developing countries typically export cash crops in exchange for manufactured goods, and fulfill their subsistent needs primarily from domestic sources [Gollin, Parente and Rogerson (2007) and Tombe (2015)].

The government's objective function is utilitarian. It has access to three Ramsey (linear) taxes: valued added tax (VAT, τ^a) on food and manufacturing goods, personal income tax (PIT, τ^w) on

[11] We argue that this assumption is innocuous in our context, since according to a large literature in labor and macroeconomics, migration in developing countries is usually driven by factors other than taxes, especially for small and moderate changes of tax rates around the status quo level. See Lucas (1997) for a review of the early work, and Lagakos, Mobarak and Waugh (2017) for recent evidence. Put differently, what matters here is not whether there is large migration *per se*, but whether migration responds to local changes of tax rates.

[12] For more details on the organization of agricultural production in low-income countries, see the handbook chapter by Eastwood, Lipton and Newell (2010).

[13] The rest of the sectors are assumed to be risk free for simplicity. For the main results to hold, we only need their risks to be quantitatively smaller. Section I.C provides more explanation on the modeling assumption.

[14] Alternatively, we can assume that there is a set of perfectly competitive trade intermediaries that trade with the exporters and households anonymously in the same way financial sectors are treated in models with default. See for instance, Azzimonti and Yared (2019 forthcoming).

households' income from formal markets, and corporate income tax (CIT, τ^r) on manufacturing firms and large farmers. We assume that the government can tax domestic agricultural sector following the evidence presented in Anderson, Rausser and Swinnen (2013) and Adamopoulos and Restuccia (2014). The government spends expenditure G on manufacturing goods. In the benchmark exercises, we assume that the expenditure is not directly valued by households. But later we allow the government to do lump-sum transfers. Such transfers can also be interpreted as an approximation of the scenario where households derive utility directly from public expenditure. We assume that the government runs balanced budget in every period, which implies that government expenditure varies between equilibria since we study different revenue mobilization scenarios.

B. *Preference*

We assume that both types of households and the large farmer share the same preference over sequences of consumption on food, manufacturing goods, and services $\mathbf{c}_t = [c_t^a, c_t^m, c_t^s]$:

$$U = \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(\mathbf{c}_t)\right],$$

where β is the time preference. We assume that the per period utility is log-linear in the three goods:

$$u(\mathbf{c}_t) = \log c_t^a + \gamma \log c_t^m + \psi \log c_t^s, \tag{1}$$

where γ and ψ are relative preferences over manufacturing goods and services, and subscript t refers to time periods.[15] With log linear preference, if the household's total consumption expenditure is $\overline{C}_t$, in the optimal consumption bundle, the household distributes the budget to c_t^a, c_t^m and c_t^s by shares $1/(1+\gamma+\psi), \gamma/(1+\gamma+\psi)$ and $\psi/(1+\gamma+\psi)$ respectively, which, given taxes and prices, further leads to

$$u(\mathbf{c}_t) = u(\overline{C}_t) = \overline{U}_t + (1+\gamma+\psi)\log \overline{C}_t, \tag{2}$$

[15]Notice that here we use the conventional homothetic utility function as opposed to the non-homothetic Stone-Geary preference with subsistence requirement $\overline{a}$ that is typically used in the development literature [Restuccia, Yang and Zhu (2008), Lagakos and Waugh (2013), and Adamopoulos and Restuccia (2014)]. We choose to do this because with homothetic preference, we can back out the consumption equivalence of welfare changes from the value functions directly, while with non-homothetic preferences, we have to rely on Monte Carlo simulations. In the literature, the Stone-Geary preference is usually used to generate a large agricultural sector in the spirit of Schultz (1953)'s "food problem," or more recently to increase relative risk aversion of the poor people [Chetty and Szeidl (2007) and Donovan (2018)]. Neither factor is crucial in our setting for several reasons. First, the rural population and hence agricultural employment in the model is exogenously given because we preclude cross-region migration. Also since we do not conduct cross-country comparison in our paper, the size of agricultural sector in our model does not have to show substantial responsiveness to changes in aggregate productivity. As a result, our model faces less challenge in generating a large agricultural sector. Second, in the earlier IMF Working Paper version of the paper [Peralta-Alva et al. (2018)], we solve a version of the model with Stone-Geary preference with $\overline{a}$ calibrated to match the food consumption share of people in the bottom quintile of the consumption distribution, but found that both the value and the impact of $\overline{a}$ are small. This alleviates the second concern.

where

$$\overline{U}_t = \log\left(\frac{1}{1+\gamma+\psi}\cdot\frac{1}{(1+\tau_t^a)p_t^a}\right) + \gamma\log\left(\frac{\gamma}{1+\gamma+\psi}\cdot\frac{1}{1+\tau_t^a}\right) + \psi\log\left(\frac{\psi}{1+\gamma+\psi}\cdot\frac{1}{p_t^s}\right),$$

is a time varying constant. As a result, the consumption-saving decision and optimal consumption bundle decision of the households can be analyzed independently. We use this property in later sections to simplify the notation.

C. *Technologies*

Informal Markets.—By working informally, rural households produce food according to the production function:

$$y_t^a = z^a \varepsilon_t^r (1-h_t^r)^{1-\alpha^a}, \tag{3}$$

where z^a is the economy-wide agricultural productivity, $1-h_t^r$ is the share of labor supplied *informally*, and $1-\alpha^a$ is the labor share. We assume decreasing returns to scale here because land transferability in developing countries is usually limited.[16] In the production function (3), we use the idiosyncratic shock ε_t^r to capture the fact that agricultural production is risky. This is especially the case in developing countries which usually lack irrigation system or hydraulic infrastructure to smooth weather shocks, or pest control services to resist outbreak of pest infestation.[17] It also captures unobserved variations in crop yields due to difference in individual farmers' ability or crop choice. Similarly, by working informally, urban households provide services according to the production function:

$$y_t^s = z^s (1-h_t^u)^{1-\alpha^s},$$

where z^s is the economy-wide services productivity, $1-h_t^u$ is the share of labor supplied *informally*, and α^s is the labor share.

Formal Markets.—We assume that labor productivity in the urban formal market is also subject to idiosyncratic shock ε_t^u. The shock reflects factors such as variations in the labor efficiency caused by matching or fluctuations in employment status. Thus the actual pre-tax labor income per unit labor in the urban area is $w^m \varepsilon_t^u$. The manufacturing firm produces by the production function:

$$y_t^m = z^m (k_t^m)^{\alpha^m} (h_t^m)^{1-\alpha^m},$$

where z^m is the manufacturing productivity, k_t^m is the total capital, α^m is the capital share, and h_t^m is the total *effective* labor units hired. Meanwhile, large farmers produce food following the same production function as (3):

$$y_t^{a,f} = z^a (h_t^a)^{1-\alpha^a}.$$

[16] For Ethiopia in particular, see the evidence presented in Chen, Restuccia and Santaeulàlia-Llopis (2017). Wu et al. (2018) provide a discussion of the land and migration policies in China. For land policies in other developing countries in general, please refer to the online appendix of Adamopoulos and Restuccia (2014).

[17] See for example, Duflo and Pande (2007), Allen and Atkin (2017), Burgess et al. (2017), Donovan (2018), and Lagakos, Mobarak and Waugh (2017).

Recall that we assume that the production of cash crops is modernized. This means that it requires modern farm machinery k_t^f:

$$y_t^* = z^*(k_t^f)^{\alpha_1^*}(h_t^*)^{\alpha_2^*},$$

where z^* is the exporting sector productivity and α_1^* and α_2^* are factor shares. For the same reason of limited land transferability before, here we assume $\alpha_1^* + \alpha_2^* < 1$. Because agricultural modernization helps greatly to reduce the risk, here for simplicity, we assume that that there is no idiosyncratic risk associated with the rural formal market.[18]

D. *Optimization Problems*

To simplify the notation, we let

$$\mathbf{C}_t^j = (1+\tau^a)(p^a c_t^{a,j} + c_t^{m,j}) + p^s c_t^{s,j}, \quad j \in \{u, r, f\},$$

denote the total consumption expenditure for urban, rural households, and large farmers.[19] Throughout it is understood that the $\mathbf{C}_t^j$s are always spent according to the optimal consumption bundle.

Households.—Let $b^{j'}$ represents savings. We assume that the ε^js are Markovian processes. The recursive problem of the urban households is:

$$V^u(b^u, \varepsilon^u) = \max_{\{\mathbf{C}^u, b^{u'}, h^u\}} \left\{ u(\mathbf{C}^u) + \beta \mathbb{E}[V^u(b^{u'}, \varepsilon^{u'}|\varepsilon^u)] \right\} \tag{4}$$

$$s.t.$$

$$\mathbf{C}^u + b^{u'} = (1-\tau^w)\varepsilon^u w^m h^u + p^s z^s (1-h^u)^{1-\alpha^s} + (1+r)b^u.$$

That of the rural households is

$$V^r(b^r, \varepsilon^r) = \max_{\{\mathbf{C}^r, b^{r'}, h^r\}} \left\{ u(\mathbf{C}^r) + \beta \mathbb{E}[V^r(b^{r'}, \varepsilon^{r'}|\varepsilon^r)] \right\} \tag{5}$$

$$s.t.$$

$$\mathbf{C}^r + b^{r'} = (1-\tau^w) w^f h^r + p^a z^a \varepsilon^r (1-h^r)^{1-\alpha^a} + (1+r)b^r.$$

The formulation above implies that more productive households (that is those with higher ε^js) choose to work in the manufacturing sector in the urban area and in their own plots in the rural area. This is consistent with the empirical evidence in La Porta and Shleifer (2014) and Eastwood, Lipton and Newell (2010). In particular, La Porta and Shleifer (2014) show that the formal sector is usually more productive because of human capital privilege, and Eastwood, Lipton and Newell (2010) argue that hired labor is less productive in agricultural sector as a consequence of moral hazard and high costs in monitoring.

[18] For the reduction of agricultural risks by modernization, see for instance, Duflo and Pande (2007) and Pingali (2007).

[19] Though we assume that large farmers also consume all kinds of goods, because in the quantitative exercise they only consist 3% of the total population, their impacts on consumption are small. It is quantitatively similar to model them as neoclassical firms operating solely in the rural area.

The two dynamic programming problems share very similar structure. The solutions to these two problems are policy functions on consumption $c^j(b,\varepsilon)$, saving $b^{j'}(b,\varepsilon)$, and labor supply to formal markets $h^j(b,\varepsilon)$, where $j=u,r$. The joint cumulative distribution functions of households in the rural and urban areas in the steady state are denoted respectively by $\Gamma^r(b^r,\varepsilon^r)$ and $\Gamma^u(b^u,\varepsilon^u)$.

Farmers and Firms.—The sequential problem of the large farmer is

$$\max_{\{\mathbf{C}_t^f,k_{t+1}^f,h_t^a,h_t^*\}} \sum_{t=0}^{\infty}\beta^t u(\mathbf{C}_t^f) \tag{6}$$

$$s.t.$$

$$\mathbf{C}_t^f + k_{t+1}^f = (1-\tau^r)(\pi_t^f+\pi_t^*) + (1-\delta)k_t^f + \tau^r\delta k_t^f,$$
$$\pi_t^f = p^a z^a (h_t^a)^{1-\alpha^a} - w^f h_t^a,$$
$$\pi_t^* = z^*(k_t^f)^{\alpha_1^*}(h_t^*)^{\alpha_2^*} - w^f h_t^*,$$

where CIT is collected over farmer's profits and the government also grants tax credit to capital investment.[20] The manufacturing firm's problem is

$$\max_{\{k_t^m,h_t^m\}} \left\{(1-\tau^r)z^m(k_t^m)^{\alpha^m}(h_t^m)^{1-\alpha^m} - w^m h_t^m - (r+\delta)k_t^m\right\}. \tag{7}$$

Notice that CIT is imposed on firm's revenue since neoclassical firm earns zero profits. The depreciation δ is also paid by the firm.

Government.—To specify the government budget constraint, we introduce several notations to ease the exposition. Define

$$C_t^x = \mu^u\int c_t^{x,u}d\Gamma^u(b_t^u,\varepsilon_t^u) + \mu^r\int c_t^{x,r}d\Gamma^r(b_t^r,\varepsilon_t^r) + \mu^f c_t^{x,f}, \qquad x\in\{a,m,s\},$$

the aggregate consumption of each goods,

$$H_t^u = \int \varepsilon_t^u h_t^u d\Gamma^u(b_t^u,\varepsilon_t^u), \quad H_t^r = \int h_t^r d\Gamma^r(b_t^r,\varepsilon_t^r),$$

as the total efficient units labor supply to the formal markets in urban and rural areas, and

$$y_t^m = z^m(k_t^m)^{\alpha^m}(h_t^m)^{1-\alpha^m},$$

the total revenue of domestic manufacturing firms. Then the government budget constraint is

$$G + \mu^f\tau^r\delta k_t^f = \tau^a(p^a C_t^a + C_t^m) + \mu^f\tau^r(\pi_t^f+\pi_t^*) + \tau^r y_t^m + \tau^w(\mu^u w^m H_t^u + \mu^r w^f H_t^r), \tag{8}$$

where $\mu^f\tau^f\delta k_t^f$ is the tax deduction to agricultural machinery investment.

[20] Appendix A.1 provides details of the solution to the large farmer's problem.

E. *Stationary Equilibrium*

We define formally the recursive competitive equilibrium in the steady state in this section. The equilibrium along the transition path could be defined similarly, hence is deferred to Appendix A.3.

Definition 1. (Recursive Competitive Equilibrium) *A recursive competitive equilibrium for the economy consists of equilibrium prices* $\mathbf{p} = \{p^a, p^s, w^m, w^f, r\}$*, value functions* $V^j(b^j, \varepsilon^j)$*, consumer decision rules* $\{c^{x,j}(b^j, \varepsilon^j), b^{j'}(b^j, \varepsilon^j), h^j(b^j, \varepsilon^j)\}$*, cumulative distribution functions* $\Gamma^j(b^j, \varepsilon^j)$*, farmer's decision rules* $\{c^{x,f}, k^f, h^r, h^*\}$*, where* $j \in \{u, r\}$ *and* $x \in \{a, m, x\}$*, and firm's decisions* $\{k^m, h^m\}$*, for any given policies* $\{\tau^a, \tau^r, \tau^w\}$*, such that*

(i) Given $\mathbf{p}$*,* $V^j(b^j, \varepsilon^j)$ *and* $\{c^{x,j}(b^j, \varepsilon^j), b^{j'}(b^j, \varepsilon^j), h^j(b^j, \varepsilon^j)\}$ *solve the households' optimization problems (4) and (5);*

(ii) Given $\mathbf{p}$*,* $\{c^{x,f}, k^f, h^r, h^*\}$ *solve the large farmer's optimization problem (6);*

(iii) Given $\mathbf{p}$*,* $\{k^m, h^m\}$ *solve the firms' optimization problem (7);*

(iv) (Aggregate Consistency) $\Gamma^j(b^j, \varepsilon^j)$*s are stationary distributions corresponding to the joint transition matrices* Π^j *constructed from* $b^{j'}(b^j, \varepsilon^j)$ *and the transition matrices of* ε^j*,* $j = u, r$*:*

$$\begin{aligned}\Pi^j &= \mathbf{Pr}[b_{t+1} = b', \varepsilon_{t+1} = \varepsilon' | b_t = b, \varepsilon_t = \varepsilon] \\ &= \mathbf{Pr}[b = b_{t+1}^{-1}(b', \varepsilon_t = \varepsilon)]\mathbf{Pr}[\varepsilon_{t+1} = \varepsilon' | \varepsilon_t = \varepsilon],\end{aligned}$$

where $b_{t+1}^{-1}(\cdot)$ *is the inverse function of saving* $b'(b, \varepsilon)$*, and we have suppressed the dependence of* j *for simplicity;*

(v) Government budget (8) is balanced;

(vi) Prices $\mathbf{p}$ *clear all markets:*

- *Urban Labor Market:*
$$\mu^u \int \varepsilon^u h^u d\Gamma^u(b^u, \varepsilon^u) = h^m.$$
- *Rural Labor Market:*
$$\mu^r \int h^r d\Gamma^r(b^r, \varepsilon^r) = \mu^f (h^a + h^*).$$
- *Capital Market:*
$$\mu^u \int b^{u'} d\Gamma^u(b^u, \varepsilon^u) + \mu^r \int b^{r'} d\Gamma^r(b^r, \varepsilon^r) = k^m.$$
- *Food:*
$$C^a = \mu^r \int z^a \varepsilon^r (1 - h^r)^{1-\alpha^a} d\Gamma^r(b^r, \varepsilon^r) + \mu^f z^a (h^a)^{1-\alpha^a}.$$

- *Services:*

$$C^s = \mu^u \int z^s (1 - h^u)^{1-\alpha^s} d\Gamma^u(b^u, \varepsilon^u).$$

- *Manufacturing Goods:*

$$C^m + \delta(k^m + \mu^f k^f) + G = z^m (k^m)^{\alpha^m} (h^m)^{1-\alpha^m} + \mu^f R^*,$$

where

$$R^* = z^* (k^f)^{\alpha_1^*} (h^*)^{\alpha_2^*},$$

is the revenue from the export sector.

Recall that we have assumed that manufacturing goods are the numeraire in this economy, meaning $p^m = 1$. Therefore out of the six markets clearing conditions, only five of them are independent, since all resource and budget constraints hold with equality automatically leads to manufacturing market clearing (the Walras' Law). In practice, we solve endogenously the prices for the first five markets. In addition, by including the R^* in the manufacturing goods market clearing condition, we have substituted in the balanced current account condition where the government imports manufacturing goods to clear trade surplus.

F. *A Simplified Economy*

We illustrate several key insights from our model using a simple example. We consider a static economy with no capital. Households are homogenous within each region. We assume that $\mu^r = 2/3, \mu^u = 1/3$ and $\mu^f = 0$. Further, we assume that $\alpha^s = 1/2$ and both the agricultural and manufacturing production functions are linear. With these assumptions, there is a closed-form solution of the model which implies that when VAT is taxed and spent on the same good, it has zero efficiency cost. Otherwise, it reduces output while increasing urban-rural gap. All proofs are deferred to Appendix D..

Result 1. *The urban-rural income gap is increasing in τ_a.*

The intuition here is that because the tax burden is imposed on all sectors, but the tax revenue is used to purchase goods produced only in the urban area, the government essentially redistributes income from the rural area to the urban area. Therefore, for this intuition to hold, it is crucial that different sectors are hosted by different regions and that there is a difference between the goods taxed and purchased by the government. When this is not the case, even though government consumption crowds out private consumption, there is no between region redistribution and total output is not affected. The next result states this formally.

Result 2. *If the government uses the tax revenue collected through value added tax to purchase the same good, then value added tax has zero efficiency cost.*

TABLE 1—EXOGENOUSLY CALIBRATED PARAMETERS

Variables	Parameters	Value
Discount Rate	β	0.96
Depreciation Rate	δ	0.06
Labor Share in Food Production	α^a	0.49
Labor Share in Services Production	$1-\alpha^s$	0.37
Capital Share in Manufacturing Production	α^m	0.37
Land Share in Cash Crop Production	α_1^*	0.19
Labor share in Cash Crop Production	α_2^*	0.32
Persistence of Urban Income Shocks	ρ^u	0.92
Persistence of Rural Income Shocks	ρ^r	0.92
Urban Population Share	μ^u	0.28
Rural Population Share	μ^r	0.69
Large Farmer Share	μ^f	0.03

III. Calibration

The model is calibrated to Ethiopia of year 2011, and the model period is one year. Broadly speaking, there are two groups of parameters to be calibrated. In the first group are those parameters whose values we take exogenously from the literature. These parameters are mostly related to the functional form of preference and technologies. The second group contains those parameters that we adjust endogenously such that certain moments implied by the model are consistent with those in the data. We explain how their values are determined one by one.

Exogenous Parameters.—We set the discount rate β and depreciation rate δ to the standard values used in the literature, 0.96 and 0.06 respectively. We assume that the factor shares of cash crops production in our model are the same as those of the agricultural sector in Adamopoulos and Restuccia (2014), leading to 0.19 and 0.32 for α_1^* and α_2^* respectively. Because food production does not use capital in our model, we assign the previous capital share to labor share, which gives us a combined labor share (which is essentially the residual from land) of 0.51. We assume that the manufacturing production function is the same as the one commonly used for the United States, giving us an α^m of 0.37. We further assume that labor share in the services production is the same as in the manufacturing sector. We note that since in our model, services are assumed to be self-employed, a different labor share would only lead to a different calibrated value of the relative sectoral productivity level, leaving the main results intact. The same argument goes for the production functions in the agricultural sector. We get the urban and rural population shares 0.28 and 0.69 from the World Bank. The 3% share of large farmer is a modeling choice. The results of the paper are affected as long as this number is small.[21] The calibration of the income shocks processes will be explained along with the endogenous parameters. Table 1 lists the values used for the exogenously calibrated parameters.

[21] As mentioned previously, an alternative choice would be to model the exporting firms as mirroring the zero-profits neoclassical manufacturing firms.

Endogenous Parameters.—The remaining parameters are calibrated jointly such that the relative consumption and production shares of different goods, the tax structure, and regional consumption Gini coefficients implied by the model are consistent with those in the data. Though varying the value of a single parameter usually leads to changes in all moments in Table 2 due to general equilibrium effects, some moments are indeed more responsive to certain parameters than to others. The matching between parameters and data moments in Table 2 highlights this relation.

We normalize the productivity of the services sector z^s to 1. We calibrate the sectoral productivity levels z^a, z^m and z^* to match the sectoral shares in output reported by National Bank of Ethiopia. The relative shares for the agricultural, manufacturing, and export sectors are respectively 42%, 33%, and 8%. In calculating these statistics, we map the manufacturing sector in our model to the data by including modern services industries such as telecommunication, banking, etc., as they resemble manufacturing firms operating on the formal market in our model. The relative preferences γ and ψ are calibrated to match the average relative consumption shares of different goods found in the *Household Consumption Expenditure Survey*. The tax structure is taken from the World Bank, where all taxes are aggregated into the three types of taxes in the model.[22]

The calibration of the income processes warrants more explanation. We assume that the two ε_t^j are both AR(1) process:

$$\varepsilon_{t+1}^j = \rho^j \varepsilon_t^j + \eta_{t+1}^j, \quad j = u, r,$$

with autocorrelations ρ^j and variances of the innovations σ_js. The data requirements for estimating these processes are very demanding.[23] Such data are rare even in developed countries, and unfortunately Ethiopia is no exception. As a result, we make some additional ad hoc assumptions on these processes, and examine the robustness of our results to these assumptions. As is stated in footnote 2, we were able to verify that the main results are indeed robust to a vast range of parameterizations thanks to many country teams in the IMF. In practice, we assume that both processes have a ρ of 0.9, the value used by Domeij and Heathcote (2004). We then calibrate σ_u^2 and σ_r^2 to match the observed consumption Ginis in the rural (0.26) and urban areas (0.40), both retrieved from the World Bank. The processes are then approximated by discrete Markov Chains using the standard Tauchen (1986)'s method. We admit that the choice of ρ is somewhat arbitrary. But since our paper focuses on the cross-sectional aspects of the economy, quantitatively, the impact of ρ is actually small. In a number of tedious sensitivity analyses, we set ρ to different levels ranging from 0.80 to 0.95, and for each ρ, recalibrate σ_u^2 and σ_r^2 again to match the consumption Ginis. We find that our results are not affected.

Table 2 lists the values used for the endogenously calibrated parameters and the model fit. Overall, the model does a reasonable job to fit the data, except that the shares of manufacturing goods

[22]Notice that because in reality, taxes take a variety of forms including exemptions and deductions, hence the calibrated tax rates are effective rates instead of statutory rates.

[23]Usually, household-level panel data with a relatively long time coverage are needed. A prominent example would be the *Panel Study of Income Dynamics (PSID)* in the United States. See for example, Storesletten, Telmer and Yaron (2004) and Meghir and Pistaferri (2011).

TABLE 2—ENDOGENOUSLY CALIBRATED PARAMETERS

Data Targets	Parameters	Value	Data	Model
Manufacturing Share in Consumption	γ	0.82	0.33	0.35
Services Share in Consumption	ψ	0.50	0.21	0.22
Rural Consumption Gini	σ_r^2	0.23	0.26	0.26
Urban Consumption Gini	σ_u^2	0.63	0.40	0.40
Tax to GDP Ratio	τ^a	0.06	0.08	0.08
CIT in Total Tax Revenues	τ^r	0.12	0.30	0.30
PIT in Total Tax Revenues	τ^w	0.06	0.17	0.19
Food Share in Output	z^a	0.39	0.42	0.34
Manufacturing Share in Output	z^m	10.19	0.33	0.38
Export Share in Output	z^*	14.17	0.08	0.10

and cash crops are slightly overshot, while food share is undershot.[24]

IV. Quantitative Results

We use the calibrated model to compare the welfare costs of revenue mobilization using different tax instruments. In the experiments, we let the government raise additional tax revenue of 2% GDP in the benchmark steady state through VAT, CIT and PIT. In all three experiments, the tax revenue is assumed to be used as government expenditure not valued directly by households. We first compute the consumption equivalence change of welfare for each tax reform and decompose the welfare changes into aggregate and distributional components following Domeij and Heathcote (2004). We then contrast the welfare costs from long-run steady state comparison and those when transitional dynamics are considered. In Section III.C, we shut down the idiosyncratic shocks in the model to isolate the impacts from risks. The results of lump-sum transfers are presented in Section III.D. We briefly talk about the policy implications from our exercises in Section III.E. The choice of 2% GDP is again arbitrary, however, in another group of sensitivity analysis, we find that the results do not change if we consider the case of 5% GDP additional tax revenue.

The quantitative exercises in this paper are different from those in search of an optimal taxation regime usually done in the literature. We do this for three reasons.

1. *Practical relevance*. Fiscal consolidations of low-income countries in the real world usually take the "additive" form considered in this paper, as opposed to more drastic redistribution of the tax responsibilities.

[24]The main reason that we are not able to get a good fit for the agricultural sector share in output is because in the model we do not allow people to migrate. Notice that since p^a enters both sides of the budget constraint of rural households, p^a is pinned down by the urban households in equilibrium. Without migration, z^a has no impact on the *demand curve for food* by urban households. As a result, put in the language of standard demand and supply curves, an increase in z^a works as a an outward shift of the supply curve with the demand curve untouched. This leads to an increase in equilibrium quantity and decrease in equilibrium price. Numerically, we find that in the equilibrium, the change in price exactly offsets that in quantity, which is a consequence of the unit elasticity implied by our log-utility assumption.

2. *Thin tax base*. Unlike advanced economies, the tax base of certain sources (for instance capital and formal wage) may be very thin in low-income countries. Hence certain tax source alone may not be enough to finance the entire public expenditure even when expropriatory tax rates are imposed.

3. *Data limitation.* High quality household-level data are rare in low-income countries. Though for the purpose of our paper—that is to investigate the trade-off of different tax instruments in an environment resembling low-income countries—the results are fairly robust across different parameterizations, optimal tax rate is likely to be heavily influenced by actual income dynamics and wealth distributions [Domeij and Heathcote (2004)]. In the absence of high quality micro-level data, we suspect that the findings from an optimal taxation exercise will be less informative.

As a result, we choose to slightly deviate from what is commonly done in the literature, and leave the discussion of optimal taxation in the context of low-income countries as an important direction for future research.

A. *Welfare Costs in the Long-run*

We first study the long-run implications of fiscal consolidations by comparing the statistics from the old and new steady states. In the benchmark calibration, the effective tax rate of VAT, CIT and PIT are respectively 6.45%, 11.55% and 5.55%. In the new steady states, these tax rates increase respectively to 10.5%, 24.0% and 15.2%. The percentage increase with respect to the benchmark rates are 62%, 108% and 174%. The variation in the percentage change here reflects difference in the size of the tax base, which reiterates our concern that certain tax base may be particularly thin in low-income countries.

Overall Welfare Costs.—In Figure 1, we show the percentage change from the benchmark steady state of six moments—aggregate output, consumption and investment, as well as regional and overall consumption Ginis—caused by the three tax instruments. As expected, the efficiency cost measured by drops in macro aggregates is the lowest for VAT. Specifically, VAT causes moderate declines in aggregate output (1.49%) and consumption (4.66%), while slightly boosts aggregate investment (1.16%) because of extra demand on manufacturing goods from the government. On the other hand, the output and consumption losses from CIT (4.85% and 7.55%) and PIT (6.19% and 9.50%) are much larger. Moreover, now the extra demand on manufacturing goods is not large enough to turn aggregate investment from falling to rising. Hence in both cases, aggregate capital in the economy also decreases (7.15% for CIT and 5.51% for PIT). The drop in investment for CIT is larger here because it directly distorts the first order condition of capital for manufacturing firms. The finding that VAT has the least efficiency cost echoes the standard Ramsey (1927) results of

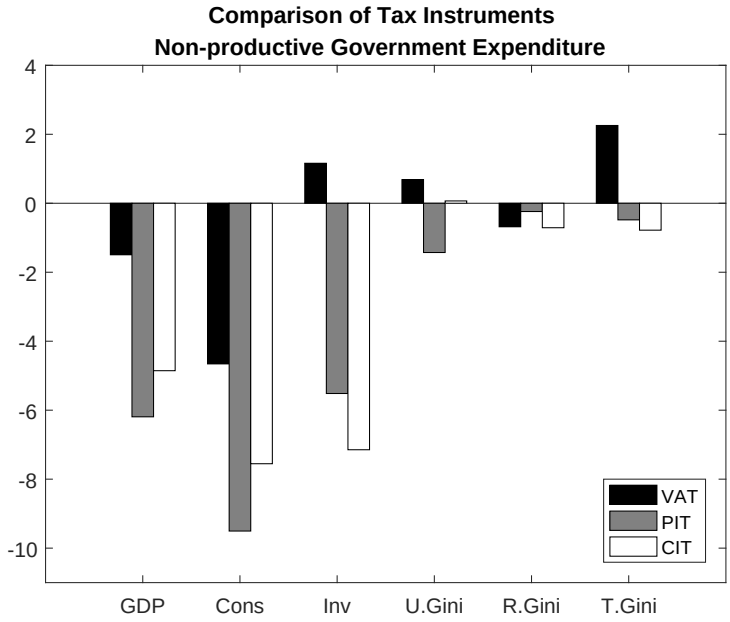

FIGURE 1. TAXES AND NON-PRODUCTIVE GOVERNMENT EXPENDITURE

Note: All results are percentage change with respect to the benchmark case.

equal taxes on all goods.[25]

Because fiscal consolidations through VAT requires the lowest percentage increase in effective tax rates, and has the least efficiency costs, it may seem that it should be preferred to the other two taxes. The welfare costs measured in consumption equivalence changes reported in Table 3, however, present a different picture. Focusing first on the last column, which shows the overall welfare costs to the economy as a whole, we find that the welfare costs of VAT (3.89%) are in fact larger than those of CIT (2.24%) and PIT (3.31%). This means that an average households would like to give up almost 4% of its annual consumption *permanently* if the government does not raise VAT rate. These costs are large. As a benchmark, the welfare costs of business cycle in the United States were estimated to be lower than 1% permanent drop of annual consumption.[26]

Regional Welfare Costs.—To see the reason of why VAT causes large welfare loss, we decompose the overall welfare costs into aggregate and distributional components following Domeij and Heathcote (2004). We find that it is the large distributional cost of VAT (1.32%) that makes it more

[25]Notice that like in many neoclassical models, here taxation leads to output loss, since most of the time they are distortive. There are exceptions of course. For instance, Anagnostopoulos and Li (2013) showed that when the utility function has 1) a constant elasticity of intertemporal substitution in consumption; and 2) a marginal rate of substitution between consumption and leisure that is proportional to consumption, then a Ramsey consumption tax is not distortive. It is, however, worth mentioning that in our model, although VAT (consumption tax) leads to increase in capital, but because it distorts allocations especially labor in other sectors, output still shrinks.

[26]The original estimates by Lucas (1987) is as low as 0.13%, while later Storesletten, Telmer and Yaron (2001) increase the estimates to around 0.8%.

TABLE 3—THE WELFARE COSTS OF FISCAL CONSOLIDATIONS

Taxes	Urban	Rural	Whole
VAT			
Total	−0.68%	−5.17%	−3.89%
Aggregate	−0.29%	−5.26%	−2.61%
Distributional	−0.46%	−0.10%	−1.32%
CIT			
Total	−2.80%	−2.02%	−2.24%
Aggregate	−2.76%	−2.25%	−2.52%
Distributional	−0.04%	0.24%	0.28%
PIT			
Total	−3.77%	−3.13%	−3.31%
Aggregate	−4.65%	−3.14%	−3.95%
Distributional	0.92%	0.02%	0.66%

costly comparing to the other two instruments. In fact, CIT and PIT all bring distributional *gains* (albeit small) as opposed to costs, because of their progressiveness. Specifically, CIT is mostly levied on capital owners while PIT on urban formal wage earners, both tend to be richer and show up at the high end of the consumption distribution.

Closer inspection reveals that the distributional cost of VAT is because the tax incidence fall mostly on rural households. In particular, while the welfare cost for urban households is only 0.68%, it is 5.17% for rural households, almost seven times *larger*. This explains that though the distributional costs are small within each region, it is of a magnitude larger if the economy as a whole is considered, meaning that it is between rather than within region redistribution that drives the results. This also can be seen from the fact that the overall consumption Gini coefficient increases by 2.3% in the case of VAT. The distinction between with and between region redistribution is important. While it is not new from the literature that consumption tax (VAT here) is often argued against in political debate as being regressive, the fundamental mechanism at work is different. In standard models, the regressiveness operates mainly through difference in the fraction of income devoted to consumption between rich and poor households, here the impact passes through cross region redistribution.

On the other hand, the tax incidence are much more evenly distributed across regions for CIT and PIT, with the urban households carry moderately higher burden than the rural households. This explains the distributional gains found in the two reforms, which can also be seen from the slight decrease in overall consumption Ginis (0.78% and 0.48% respectively). The reason that VAT has unevenly distributed tax burdens is exactly the implicit income redistribution driven by the difference in entities that the taxes are collected and spent demonstrated in Results 1 and 2.

Aggregate and Distributional Components.—Despite that the distributional components are crucial to elevating the welfare costs of VAT relative to CIT and PIT, most of the total welfare costs are driven by the aggregate components.

For CIT and PIT, because the additional tax burden falls mostly on a small number of rich people, their distributional impacts are small. This can be indirectly inferred from the fact that the

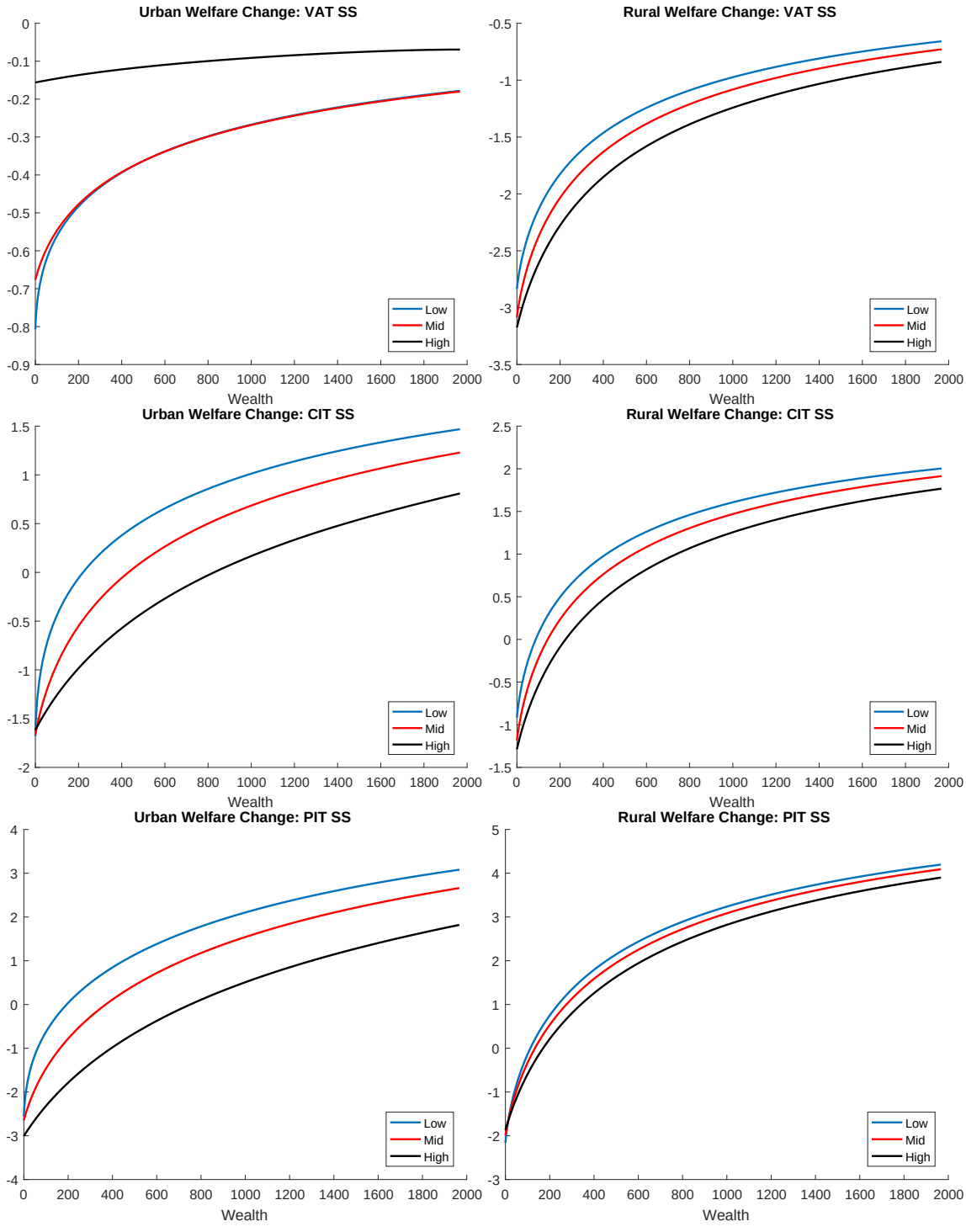

FIGURE 2. WELFARE COSTS ACROSS WEALTH AND INCOME LEVELS

percentage increase in effective tax rate is much larger for these two scenarios. In Figure 2, we plot the changes in value function for households with different productivity levels and savings. The cases for CIT and PIT can be seen from the lower four panels, where households with higher productivity are subject to higher welfare loss. Given individual productivity level, households with more savings are less affected by fiscal consolidations, suggesting their better ability to insure consumption against shocks to disposable income. However, as is shown in the top-left panel of Figure 2, within the urban area, VAT is regressive. This is because the increase in aggregate investment pulls up the wage for formal workers, who are more likely to be richer.[27]

In addition, although VAT causes the least decline in aggregate consumption, the consumption equivalence costs from the aggregate components are in fact higher. This suggests that distortions to the optimal consumption bundle due to first order impacts on the relative prices of goods and services by VAT have sizable welfare consequences.[28] However, despite this, the informality problem is not worsened by any of the taxes for two reasons. The first is the decrease in household disposable income which subsequently suppresses the demand for services; and second is the extra demand on manufacturing goods from the government. Specifically, labor supply to the informal sector only increases by 1.4% and 0.16% with VAT and PIT reforms, while for CIT reform, it actually decreases by 1.9%.

B. *Transitional Dynamics*

The formal definition of the equilibria along the transition is provided in Appendix A.3. We assume that the economy begins from the benchmark steady state. The corresponding tax rates are then unexpectedly and permanently increased to the levels in the new steady states. The economy is then simulated forward until convergence to the new steady state. Table 4 compares the welfare costs calculated from steady-state comparisons with those when transitional dynamics are included for rural and urban areas.

We find that unlike previously found in the literature, the difference between long-run and short-run results are small in our context. This is because in low-income countries, capital stock and hence household savings are low. Changes of these variables induced by the tax reforms considered here are also small. Because savings and capital are the state variables that control the intertemporal decisions of agents in the model, their quick adjustment implies that the transition to the new steady

[27] The logic here is similar to the one that Dávila et al. (2012) use to study how factor prices and the wealth distribution jointly determines whether an economy is constrained efficient.

[28] Importantly, this is not completely driven by the fact that aggregate consumption decreases more in the rural area, and hence from the perspective of a utilitarian government the welfare costs are larger. Specifically, rural aggregate consumption drops by 7.3% and 6.9% for the VAT and CIT respectively. However, the aggregate component of VAT almost doubles that of CIT.

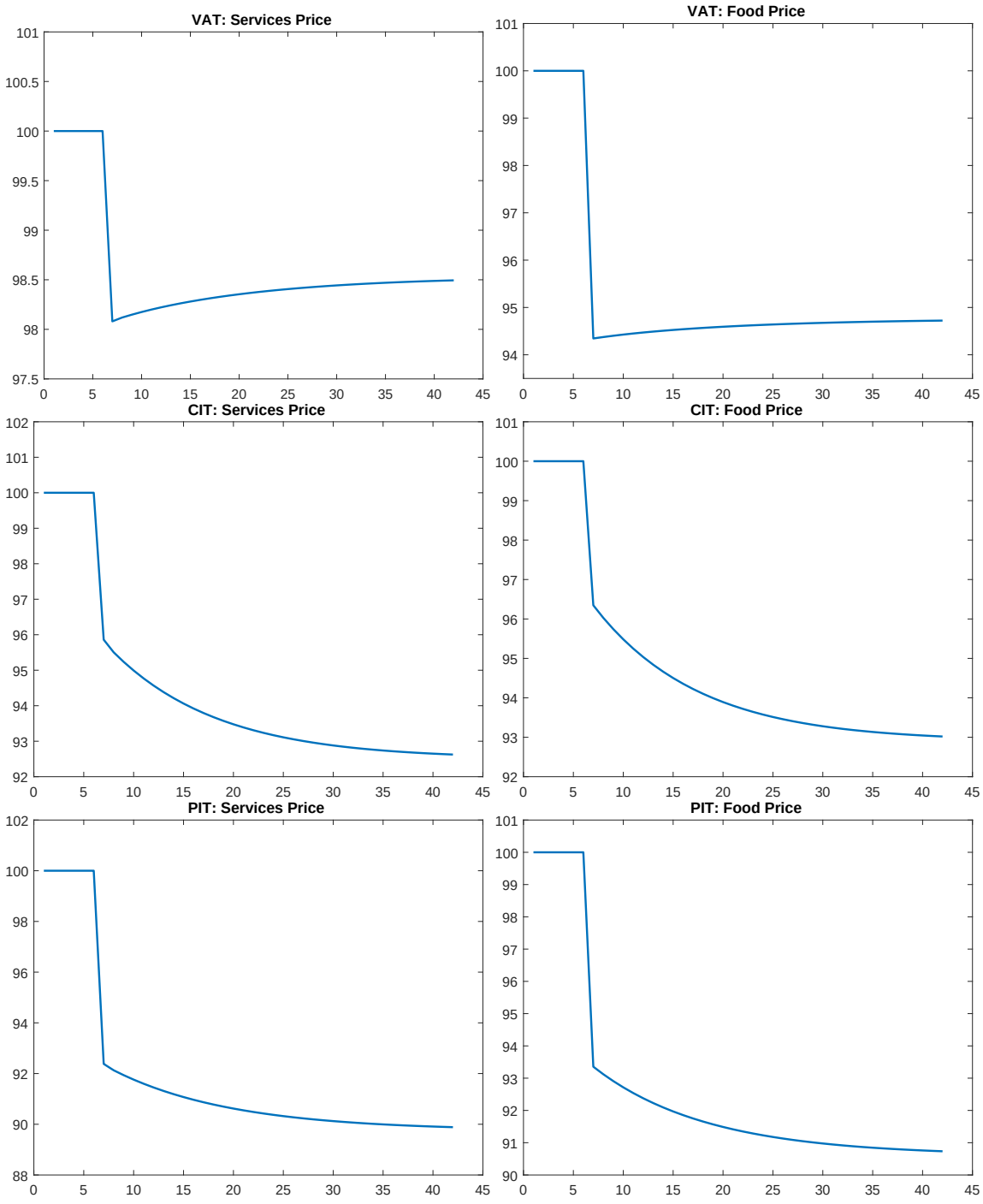

FIGURE 3. TRANSITION PATHS OF FOOD AND SERVICES PRICES

TABLE 4—SHORT-RUN VERSUS LONG-RUN WELFARE COSTS

	Urban			Rural		
	Total	Aggregate	Distributional	Total	Aggregate	Distributional
VAT						
Steady State	−0.68%	−0.29%	−0.39%	−5.17%	−5.26%	0.10%
Transition	−0.95%	−0.50%	−0.46%	−5.22%	−4.96%	−0.27%
CIT						
Steady State	−2.80%	−2.76%	−0.04%	−2.02%	−2.25%	0.24%
Transition	−2.17%	−2.09%	−0.08%	−1.85%	−1.72%	−0.13%
PIT						
Steady State	−3.77%	−4.65%	0.92%	−3.13%	−3.14%	0.02%
Transition	−3.33%	−3.86%	0.55%	−3.16%	−2.82%	−0.36%

state is also fast.[29] For this reason, out of the three tax reforms, the transition of CIT and PIT take longer periods because capital stock decreases by more, yielding larger difference between long-run and short-run effects compared to the case of VAT. Similarly, across the two regions, since the average level of saving in the rural area is only about 7% of the average level in the urban area, the convergence of rural area is almost immediate due to this extremely low level of saving, making the costs between long-run and short-run nearly negligible.

Figure 3, which contains the transition paths of prices for food and services for the three tax reforms, further makes the point. Here we only plot the transition paths for the two consumption goods because when savings are low, on average, the welfare of households is more responsive to prices as opposed to interest rate changes. Furthermore, the total welfare costs when short-run dynamics are considered are smaller for CIT (1.94% versus 2.24%) and PIT (3.21% versus 3.31%), but larger for VAT (4.01% versus 3.89%), because capital stock in the new steady state is lower for CIT and PIT reform, allowing the households to consume more during the transition.

C. *The Role of Idiosyncratic Risks*

Development economics studies have shown that uninsurable idiosyncratic risks have profound impacts on the welfare of households.[30] In this section, we investigate how these risks shape the welfare costs associated with fiscal consolidations. To do this, we shut down the idiosyncratic

[29]The number of periods it takes to converge to the new steady state should not be interpreted too literally though, since our model is really disciplined only by the cross-sectional properties of the real economy. Therefore, for example, the middle panels of Figure 3 should *not* be interpreted as it would take around 40 years for CIT reform to converge. If we wish to push the model harder on its time dimension, we would have to calibrate parameters that pin down the intertemporal decisions of the households more carefully, for instance the discount factor β, depreciation rate δ, and the autocorrelation ρ. We do not push the model hard here as a compromise of tractability of the overall calibration exercise and data limitations.

[30]Recent studies discuss how these risks affect migration, risk sharing [Lagakos, Mobarak and Waugh (2017), Morten (2019)], and agricultural productivity [Donovan (2018)].

TABLE 5—THE WELFARE COSTS WITH AND WITHOUT IDIOSYNCRATIC RISKS

	VAT		CIT		PIT	
	Benchmark	No Risk	Benchmark	No Risk	Benchmark	No Risk
Total	−4.01%	−0.59%	−1.94%	−0.33%	−3.21%	1.91%
Aggregate	−2.58%	−2.55%	−1.91%	−2.21%	−3.37%	−2.74%
Distributional	−1.46%	2.01%	−0.02%	1.92%	0.17%	4.77%

shocks by assuming that $\varepsilon^u = \varepsilon^r = 1$ while keeping all the other parameters untouched.[31] Naturally, all the prices and allocations in both the stationary equilibria and during the transitions will be different from the Ethiopian economy. Therefore, the results in this section should be interpreted as what would happen if suddenly the idiosyncratic risks diminish, as opposed to approximating the Ethiopian economy with a representative agents model. In absence of idiosyncratic risks, the Euler equations of households reduce to

$$\frac{1}{\beta} = 1 + r,$$

implying that the interest rate is constant. This further leads to constant wage as well. However, interest rate in fact varies a lot during the transition. Therefore for all the results in this section, we include transitional dynamics in the computation.

The first difference we find when contrasting the steady state equilibria with and without risks under the benchmark calibration is that when risks are eliminated, capital stock and hence savings increase by 47%. At its face value, it seems to suggest an absence of "precautionary saving" in our model. This finding connects to the literature showing that savings may indeed be less when labor supply is endogenous, but for a different mechanism. For instance, in Marcet, Obiols-Homs and Weil (2007), wealth effect that discourages labor supply is the key mechanism at work. Here it is because that when there are idiosyncratic risks, households supply less labor to the formal sector to reduce their exposure to income fluctuations. In this sense, the mechanism shows more of a resemblance to the lower usage of intermediate goods caused by risk aversion in Donovan (2018).

Because of this and the fact that the model features aggregation [Chatterjee (1994) and Azzimonti, de Francisco and Krusell (2008)], the wealth distribution in the no risk case is indeterminate and will be different from that in the benchmark equilibrium. As a result, when migrating the wealth distribution in the benchmark economy to the no risk case, we scale up the distribution such that the aggregate saving equals that in the no risk equilibrium. Table 5 compares the welfare costs of the models with and without risks when the economy as a whole is considered. We find that overall, the aggregate components when the risks are shut down are similar, with the distributional components showing substantial discrepancy. Similar to the benchmark case, VAT reform here also leads to the lowest consumption decline, which is 5.17% compared to 7.49% and 8.78% for

[31] Notice that shutting down the idiosyncratic risks is different from completing the markets, though the market in the representative agent model is indeed complete. Completion of the financial markets usually leads to higher welfare gains since it allows households to take advantage of the positive productivity shocks. See Pijoan-Mas (2006) for more details.

CIT and PIT reforms. But again, at the same time, the gaps in aggregate components between VAT and the other two taxes are much less than the those in aggregate consumption. This once again indicates that the distortion in relative prices by VAT has first order welfare consequences in our model. In addition, because the aggregate component is essentially the total welfare costs in pure representative agent model, the fact that the aggregate component is the lowest for CIT—the counterpart of capital income tax in our model—suggests that the standard Chamley-Judd result of zero capital income tax may be different in the context of low-income country economies. Since optimal taxation is not the main subject of this paper for the reasons elaborated at the beginning of this section, we leave this for future research.

Across all tax reforms, we find that idiosyncratic risks lead to large distributional costs of fiscal consolidations. In fact, from the perspective of redistribution, all tax reforms now bring large welfare *gains* instead of losses. This happens because in the no risk case, on average households are wealthier. Hence savings weight more than per period income in their total disposable income. Thus for the utilitarian government, the drops in goods prices in the tax reforms benefit poor people more because of their higher marginal utility. Eventually, this leads to more evenly distributed indirect utility. On the contrary, in the benchmark case where the levels of savings are low, changes in goods prices affect simultaneously the income and expenditure of the households in quantitatively significant manner, leaving the distributional components small.

D. *Lump-sum Transfers*

The results so far suggest that the major disadvantage of VAT is its unevenly distributed tax burdens while those for CIT and PIT are their efficiency costs. It is natural to ask if fiscal consolidation by VAT is paired with lump-sum transfers to the rural area or similarly if CIT and PIT are accompanied with pro-growth policy, to what extent would the welfare costs be mitigated? This section answers this question using the VAT-Transfer combination as illustration. We choose to focus on VAT also because usually it has the largest tax base in low-income countries and for this particular reason is more prevalent. Furthermore, since in Section III.B, we have shown that transitions are in general fast for the benchmark calibration, in this section we focus mostly on the long-run effects except for Case 1 below. Specifically, we consider three experiments in this section.

- Case 1: Half of the additional tax revenue is used for lump-sum transfers to rural households.
- Case 2: Half of the additional tax revenue is used for lump-sum transfers to the whole population.
- Case 3: All additional tax revenue is used for lump-sum transfers to the whole population.

Notice that Case 3 can also be interpreted as households taking utility directly from additional government expenditure.

The long-run and short-run welfare costs for Case 1 is summarized in Table 6. Contrasting the results in Tables 3 and 6, we find that overall, about 66.8% of the welfare costs from the VAT reform

TABLE 6—THE WELFARE COSTS WITH LUMP-SUM TRANSFERS TO RURAL HOUSEHOLDS

	Urban	Rural	Whole
Transition			
Total	−1.43%	−1.29%	−1.33%
Aggregate	−1.37%	−1.68%	−1.51%
Distributional	−0.07%	0.40%	0.19%
Steady State			
Total	−1.29%	−1.30%	−1.29%
Aggregate	−1.32%	−1.95%	−1.72%
Distributional	0.03%	0.66%	0.32%

is mitigated.[32] This comes from a significant improvement in the distribution of tax burdens across regions, with the overall distributional component being a 0.19% gains rather than a 1.46% loss. Within the rural area, the distributional component also turns from a 0.10% loss to 0.40% gains. Because with lump-sum transfers to the rural households, the government spends less expenditure on manufacturing goods, the welfare costs that the urban households face are larger, echoing the implicit redistribution of VAT as demonstrated in Result 1.

We then calculate the changes in macro aggregates of Cases 1 to 3 with respect to the case where all tax revenues are used to purchase manufacturing goods. The impacts of different lump-sum transfers programs are summarized in Figure 4. In each group of bars, from left (black) to right (white), we plot the results of Cases 1 to 3 as specified above.

By comparing Cases 1 and 2, we can isolate the differential responses of rural and urban households to transfers. In particular, we find that Case 1 leads to higher increase in output and consumption, and overall larger improvement in consumption inequality. This is because rural households increase their consumption more than urban households in response to cash transfers due to their higher marginal propensity of consumption [Carroll and Kimball (1996)]. The decline in aggregate investment is higher in Case 2, because urban households' income is less risky and hence the precautionary saving declines. The comparison of cases 2 and 3 illustrates the effects of a uniform expansion of cash transfer programs, with around a doubling of the impact from Case 2.

E. *Policy Implications*

The quantitative results so far draw several lessons into the design of the fiscal consolidation packages focusing specifically on low-income countries that were not extensively studied before.

First, the fact that the regressiveness of consumption taxes is driven in principle by between-region in addition to the within-region redistribution suggests that the trade-off of VAT should be more carefully weighted in low-income countries before implementing. Moreover, theoretically, the finding points to a new margin where in practice taxes can cause unintended side effects. That is whenever there is mismatch between the identity of tax incidence and government expenditure,

[32] One caveat is that because here less resources are "wasted" in the sense of not directly valued by the households, the comparison of Case 1 with the no transfer case is not a "fair" one because households do value transfers.

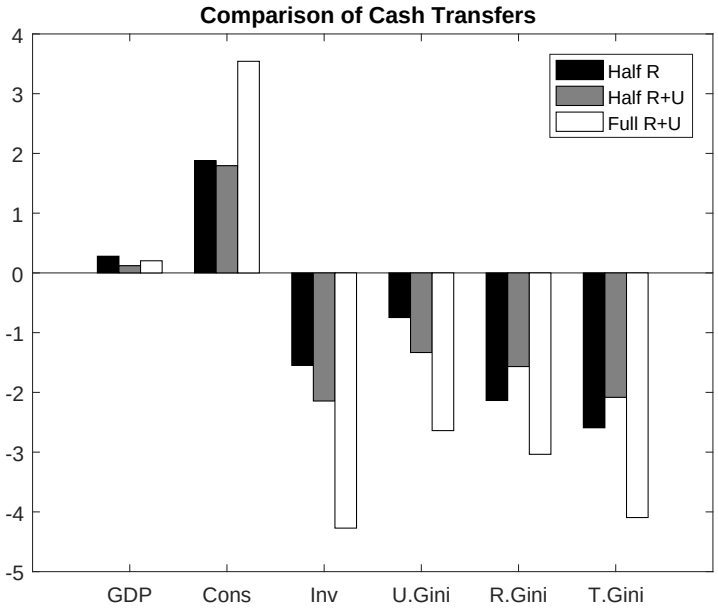

FIGURE 4. THE IMPACTS OF CASH TRANSFER PROGRAMS

Note: All results are percentage change with respect to the VAT case.

important redistributional implications are likely to follow. Second, tax instruments like CIT and PIT that were previously considered as more costly could have lower welfare costs due to their fairer distribution of tax incidence across regions. Third, because VAT also distorts households' optimal consumption bundle, decrease in aggregate consumption brings large welfare loss. Fourth, due to their low capital stock, the full long-run effect of a policy reform is predicted to be realized pretty quickly. Last, the negative impacts to the economy from fiscal consolidations are most effectively remedied by regional transfer for VAT, and pro-growth policies for the PIT and CIT.

V. Conclusion

In this paper, we quantitatively investigate the welfare costs of fiscal consolidations in low-income countries through VAT, CIT and PIT. We find that VAT has the least efficiency costs but causes large welfare costs due to the implicit redistribution of income from poorer rural to richer urban households. CIT and PIT, on the other hand, although have higher efficiency costs, but result in substantially lower welfare costs since the tax incidence are distributed more evenly between regions. We find that including short-run dynamics during the transition are less important in the context of low-income countries, because with low capital stock, the economy usually converges rapidly to the new equilibrium. We also find that idiosyncratic risks are associated with large distributional costs of taxation which are mainly driven by the difference in the composition of income in the two scenarios. Furthermore, we find that lump-sum transfers to the rural area are

able to compensate for two-thirds of the welfare costs for the VAT reform.

Our results suggest that the unique economic structure low-income countries indeed alters how different tax instruments work in nontrivial ways, which has profound implications on the proposed and implemented policy reforms in low-income countries facing challenging fiscal conditions. It also opens up several avenues for follow-up research. For instance, as we briefly touched upon in the beginning of Section III, quantifying the impacts of the canonical revenue neutral tax reforms considered in the literature for low-income countries is no doubt an important task. This also opens meaningful discussion of political support from different group of households for various reforms.[33] It is also important to investigate quantitatively other aspects of taxation (like optimal progressiveness of labor income tax) in the environment of low-income countries. Moreover, we believe that to figure out whether the classical optimal taxation results under complete markets still hold when being migrated to the environment of low-income economies has both theoretical and policy values. The results in Section III.C suggests that this may not be the case. Another direction that we believe is important is to use richer data to quantify carefully the macroeconomic and development implications of idiosyncratic risks [Lagakos, Mobarak and Waugh (2017) and Feng, Lagakos and Rauch (2018)]. We leave these extensions to future work.

[33] In our current exercises, because taxes are always distortive, all households suffer from welfare loss. Figure 2 shows the welfare changes.

ONLINE APPENDIX (NOT FOR PUBLICATION)

A. *Computational Algorithms*

This appendix lays out the algorithm to compute both the steady state equilibrium and the transition path between steady states. We start with the solution to the large farmer's problem which we skip in the main text. We then present the pseudo code to solve the steady state equilibrium in Section A.2. The definition of the transition path and the algorithm to solve it are introduced respectively in Sections A.3 and A.4.

A.1 *The Large Farmer's Problem*

For convenience, we write down the sequential problem of the larger farmer below again, but with a time dimension indicating the transition path:

$$\max_{\{\mathbf{C}_t^f, k_{t+1}^f, h_t^a, h_t^*\}} \sum_{t=0}^{\infty} \beta^t u(\mathbf{C}_t^f) \tag{A.1}$$

$$s.t.$$

$$\begin{aligned}
&\mathbf{C}_t^f + k_{t+1}^f = (1-\tau_t^r)(\pi_t^f + \pi_t^*) + (1-\delta)k_t^f + \tau_t^r \delta k_t^f, \\
&\pi_t^f = p_t^a z^a (h_t^a)^{1-\alpha^a} - w_t^f h_t^a, \\
&\pi_t^* = z^*(k_t^f)^{\alpha_1^*}(h_t^*)^{\alpha_2^*} - w_t^f h_t^*,
\end{aligned}$$

First, notice that the decision of labor used in domestic food production h_t^a is always intra-temporal. It is determined by the first order condition:

$$(1-\alpha^a) p_t^a z^a (h_t^a)^{-\alpha^a} = w_t^f, \tag{A.2}$$

where the time dimension is reflected in prices p_t^a and w_t^f. Second, the labor hired in cash crops production h_t^* is determined by the following first order condition:

$$\alpha_2^* z^* (h_t^*)^{\alpha_2^*-1}(k_t^f)^{\alpha_1^*} = w_t^f, \tag{A.3}$$

which relies on the level of capital k_t^f. Notice that with h_t^a and h_t^*, we are also able to calculate π_t^a and π_t^*. As a result, the larger farmer's problem boils down to solving for the sequence of consumption $\{\mathbf{C}_t^f\}_{t=0}^{\infty}$ and saving $\{k_{t+1}^f\}_{t=0}^{\infty}$ given the sequence of prices $\{\mathbf{p}_t\}_{t=0}^{\infty}$. As is standard in the consumption-saving problem, these sequences are pinned down by the Euler Equations and the budget constraints. By using the 1-1 relationship between c_t^a, c_t^m, c_t^s and total consumption expenditure C_t specified in Section I.B, the problem can be reduced to solving for the sequences of service consumption and saving. In particular, the Euler Equations are

$$\frac{1}{p_t^s c_t^s} = \frac{1}{p_{t+1}^s c_{t+1}^s} \beta \left[(1-\delta) + \delta\tau_{t+1}^r + (1-\tau_{t+1}^r) \frac{\partial \pi_{t+1}^*}{\partial k_{t+1}^f} \right], \tag{A.4}$$

where we have substituted in the first order conditions of c_t^s

$$\frac{1}{c_t^s} = \lambda_t p_t^s, \quad \forall t,$$

with λ_t representing the Lagrangian multipliers of the budget constraints, and the marginal product of capital

$$\frac{\partial \pi_{t+1}^*}{\partial k_{t+1}^f} = \alpha_1^* z^* (h_{t+1}^*)^{\alpha_2^*} (k_{t+1}^f)^{\alpha_1^* - 1}.$$

The budget constraints are

$$\left(\frac{1+\gamma+\psi}{\psi}\right) p_t^s c_t^s + k_{t+1}^f = (1-\tau_t^r)(\pi_t^a + \pi_t^*) + (\tau_t^r \delta + 1 - \delta) k_t^f. \tag{A.5}$$

Solution in the Steady State.—Using the fact that in the steady state, the values of all variables remain the same, we are able to simplify both the Euler Equations and the budget constraints. More specifically, the Euler Equation (A.4) now becomes

$$\frac{1}{\beta} = (1-\tau^r)\frac{\partial \pi^*}{\partial k^f} + 1 - \delta + \tau^r \delta, \tag{A.6}$$

and the budget constraint (A.5) is

$$\left(\frac{1+\gamma+\psi}{\psi}\right) p^s c^s = (1-\tau^r)(\pi^a + \pi^*) + (\tau^r \delta - \delta) k^f. \tag{A.7}$$

Substitute (A.3) into (A.6), we can directly solve for k^f. Equation (A.7) can then be used to find c^s, which completes the characterization of the solution to the large farmer's problem in the steady state.

A.2 *Computing the Steady State Equilibrium*

The model is solved by discretization. To proceed, we denote the asset space by $\mathcal{B}$, and the idiosyncratic shocks spaces as $\mathcal{E}_r, \mathcal{E}_u$ for rural and urban households. The pseudo code is given as follows.

1. Discretize the state spaces. We use i for the index of the saving grids, and j for the index of the shock grids.
   (i) Construct n geometric grids $\mathbf{B} = \{b_1, \cdots, b_n\} \subseteq \mathcal{B}$.[34]
   (ii) Approximate the shock distributions of productivity in rural informal and urban formal sectors using Tauchen (1986)'s method. Denote the m-states Markov chain by $\boldsymbol{\varepsilon}_x = \{\varepsilon_{1,x}, \cdots, \varepsilon_{m,x}\} \subseteq \mathcal{E}_x, x \in \{r, u\}$.

[34]Geometric grids puts more grids when b is small, since according to the consumption saving literature [Aiyagari (1994) and Carroll (1997)], this region is where the policy functions are likely to be nonlinear.

2. Make an initial guess for the price vector $\mathbf{P}_0 = \{p^a, p^s, w^m, w^f, r\}$.

3. Given $\mathbf{P}_0$, solve the recursive optimization problems of the urban and rural households (4) and (5). Because the two problems have analogous structure, we take the urban problem as an example here. We use the *Feasible Sequential Quadratic Programming (FSQP)* method [Zhou, Tits and Lawrence (1997)] to solve the Bellman equation. Specifically, the FSQP solves the minimization problem

$$\min\{f(x)\} \quad s.t. \quad g_j(x) \leq 0,$$

by first approximating $f(x)$ with a quadratic form and solve the resulting constrained quadratic optimization problem using first order conditions. To do this, the expected continuation value in the Bellman Equation $\mathbb{E}[V(b', \varepsilon'|\varepsilon)]$ is approximated by *Monotone Piecewise Cubic Hermite Interpolation (PCHI)* [Fritsch and Carlson (1980), Fritsch and Butland (1984)].

   (i) For each node (b_i, ε_j), calculate the optimal labor supply decision. Notice that this decision is intra-temporal and is independent from the intertemporal consumption saving decision.

   (ii) Make an initial guess on the value function at $V(b_i, \varepsilon_j)$. Our choice here is the one that household uses all resource for consumption.

   (iii) Construct the PCHI from the Lagrangian table. This includes the values for both the polynomial and its derivative at each node. In addition, we also need to calculate the derivative of the constraint with respect to the choice variables c and b'.

   (iv) Use the FSQP to solve the optimization problem at (b_i, ε_j).

   (v) Loop over all nodes in $\mathbf{B}$ and $\boldsymbol{\varepsilon}_u$ for step (ii) and (iii) to get the value function $V(b, \varepsilon)$ and policy functions $c(b, \varepsilon)$ and $b'(b, \varepsilon)$.

   (vi) Compare the difference between the initial guess of both value functions and policy functions in steps (ii) and (v). If both sets of functions converge, continue, otherwise update the guess in (ii) and repeat until (v).

4. Calculate the solution of the large farmer's problem by using Equations (A.6) and (A.7).

5. Construct the joint *density* functions of the stationary asset-shock distributions $\phi_{i,j}$ for both the urban and rural households. Again due to symmetry, we suppress the indicator for the rural and urban areas. The cumulative distribution functions of the two densities are $\Gamma^u(b_i, \varepsilon_j)$ and $\Gamma^r(b_i, \varepsilon_j)$. We use Monte Carlo simulation on the density functions to compute the the invariant joint density functions.

   (i) For dimension i of $\phi_{i,j}$, construct denser nodes on $\mathcal{B}$. In particular, we construct linear grids $\overline{\mathbf{B}}_1 = \{\bar{b}_1, \cdots, \bar{b}_{n_1}\}$, where $n_1 > n$.

   (ii) Construct an initial distribution. Here we set the marginal density on i to be the invariant distribution of the idiosyncratic shocks, and the marginal density on j to be the uniform distribution.

(iii) For agents with states (b_i, ε_j), the saving policy function $b'(b_i, \varepsilon_j)$ and the transition matrix of the idiosyncratic shock determine the share of population at $(b'(b_i, \varepsilon_j), \varepsilon'_j)$, where the Law of Large Number is implicitly invoked. Notice that $b'(b_i, \varepsilon_j)$ may not be on $\overline{\mathbf{B}}_1$. Under such scenarios, we first locate the interval in $\overline{\mathbf{B}}_1$ that contains $b'(b_i, z_j)$, we then distribute the share of people to the endpoints of the interval based on the relative distance between $b'(b_i, \varepsilon_j)$ and the endpoints. However, ε'_j will always be on $\boldsymbol{\varepsilon}_x$, owing to the discrete nature of the Markov Chain.

(iv) Repeat (iii), until the density functions from two successive iterations converge, where by stationarity the density function is that of the invariant distribution.

(v) From $\phi_{i,j}$, we can construct stationary distributions over consumption, income, and savings using the same Monte Carlo simulation method.

6. Using invariant distributions $\phi_{i,j}$, we can aggregate individual decisions to find the aggregate demand and supply. If the aggregate demand does not equal to supply, return to Step 2 and make a new guess, and iterate until demand equals supply. Here we use Powell's Hybrid Method to solve for the market clearing $\mathbf{P}$.[35]

In practice, the model is parallelized using MPI over the shock states ε_j when solving the Bellman Equations (Step 3).

A.3 *Equilibria along the Transition Path*

We assume that the model is perfect foresight along the transition path, meaning that all economic agents predict correctly the evolution of all prices and policy variables along the transition path [Ríos-Rull (1999) and Domeij and Heathcote (2004)]. For the ease of exposition, we use tax reform as example. Suppose that the tax rates in the status quo are $\boldsymbol{\tau}_o = \{\tau_o^a, \tau_o^w, \tau_o^r\}$ and those after the reform are $\boldsymbol{\tau}_n = \{\tau_n^a, \tau_n^w, \tau_n^r\}$, where the subscripts o and n are short for "old" and "new" respectively. We maintain the same notation throughout the rest of the section. The tax rates are the only difference in parameters between the two steady states. Assuming that the transition takes T periods, we will explain how T is determined shortly after. In period 1, the economy is at the old steady state and starting from period T the economy is at the new steady state. We assume that all agents hold the anticipation that all taxes stay at $\boldsymbol{\tau}_o$ until $t = 1$ (included). At $t = 2$, the tax rates change to $\boldsymbol{\tau}_n$ as a surprise and remain at the new level forever. Starting from $t = 2$, all agents in the economy correctly anticipate that the tax rates are $\boldsymbol{\tau}_n$ from then on, and moreover, the paths of the prices $\{\mathbf{p}_t\}_{t=1}^{\infty}$ are the actual realized market clearing ones.

Unlike the case of steady state, because the prices and policies are changing along the transition, all decision problems are not time invariant. Since the economy converges to the new steady state

[35] In the actual program, we use the implementation provided by the routine `hybrd` from the MINPACK. Numerically, the Powell's Hybrid Method works best when the Jacobian matrix is diagonally-dominating, meaning that changes of one price have only secondary effect on other markets. Our model does not behave in this way, since the formal and informal sectors in the urban area are closely connected, hence changes in p^s and w^m have mutually first order impacts on the services good and urban formal labor markets. As a result, it is necessary that the initial guess is within the proximity of the true zeros.

after T periods, all agents take that as given and determine their decision functions during transition by backward induction. More specifically, the dynamic programming problem for the urban households is

$$V_t^u(b_t^u, \varepsilon_t^u) = \max_{\{\mathbf{C}_t^u, b_{t+1}^u, h_t^u\}} \left\{u(\mathbf{C}_t^u) + \beta \mathbb{E}_t[V_{t+1}^u(b_{t+1}^u, \varepsilon_{t+1}^u|\varepsilon_t^u)]\right\} \tag{A.8}$$

$$s.t.$$

$$\mathbf{C}_t^u + b_{t+1}^u = (1-\tau_t^w)\varepsilon_t^u w_t^m h_t^u + p_t^s z^s (1-h_t^u)^{1-\alpha^s} + (1+r_t)b_t^u,$$

and that of the rural households is

$$V_t^r(b_t^r, \varepsilon_t^r) = \max_{\{\mathbf{C}_t^r, b_{t+1}^r, h_t^r\}} \left\{u(\mathbf{C}_t^r) + \beta \mathbb{E}_t[V_{t+1}^r(b_{t+1}^r, \varepsilon_{t+1}^r|\varepsilon_t^r)]\right\} \tag{A.9}$$

$$s.t.$$

$$\mathbf{C}_t^r + b_{t+1}^r = (1-\tau_t^w)w_t^f h_t^r + p_t^a z^a \varepsilon_t^r (1-h_t^r)^{1-\alpha^a} + (1+r_t)b_t^r,$$

where $V_T^u(\cdot)$ and $V_T^r(\cdot)$ are the value functions in the new steady state.

Similarly for the large farmers, using the fact that k_T^f is the new steady state level, Equations (A.4) and (A.5) can be used to solve $\{\mathbf{C}_t^f, k_t^f\}_{t=2}^{T-1}$ backwards.

Now we are ready to define formally the equilibria along the transition path.

Definition 2. (Perfect Foresight Competitive Equilibria during Transition) *The perfect foresight competitive equilibria during the transition for the economy consists of equilibrium prices* $\{\mathbf{p}_t\}_{t=1}^T = \{p_t^a, p_t^s, w_t^m, w_t^f, r_t\}_{t=1}^T$*, value functions* $\{V_t^j(b^j, \varepsilon^j)\}_{t=1}^T$*, consumer decision rules* $\{c_t^{x,j}(b^j, \varepsilon^j)$*,* $b_{t+1}^j(b^j, \varepsilon^j)$*,* $h_t^j(b^j, \varepsilon^j)\}_{t=1}^T$*, cumulative distribution functions* $\{\Gamma_t^j(b^j, \varepsilon^j)\}_{t=1}^T$*, farmer's decision rules* $\{c_t^{x,f}, k_{t+1}^f, h_t^r, h_t^*\}_{t=1}^T$*, where* $j \in \{u, r\}$ *and* $x \in \{a, m, x\}$*, and firm's decisions* $\{k_t^m, h_t^m\}_{t=1}^T$*, for any given policies* $\{\tau_t^a, \tau_t^r, \tau_t^w\}_{t=1}^T$*, such that*

(i) The economy is at the old steady state at $t = 1$ *and at the new steady state from* $t = T$*.*

(ii) Given $\mathbf{p}_t$*,* $\{V_t^j(b^j, \varepsilon^j)\}_{t=1}^T$ *and* $\{c_t^{x,j}(b^j, \varepsilon^j), b_{t+1}^j(b^j, \varepsilon^j), h_t^j(b^j, \varepsilon^j)\}$ *solve the households' optimization problems (*A.*8) and (*A.*9);*

(iii) Given $\mathbf{p}_t$*,* $\{c_t^{x,f}, k_{t+1}^f, h_t^r, h_t^*\}$ *solve the large farmer's optimization problem (*A.*1);*

(iv) Given $\mathbf{p}_t$*,* $\{k_t^m, h_t^m\}$ *solve the firms' optimization problem (7) at each* t*;*

(v) (Aggregate Consistency) The joint transition matrices Π_t^j *of* $\Gamma_t(b^j, \varepsilon^j)$ *are constructed from* $b_{t+1}^j(b^j, \varepsilon^j)$ *and the transition matrices of* ε^j*,* $j = u, r$ *are:*

$$\begin{aligned}\Pi_t^j &= \Pr[b_{t+1} = b', \varepsilon_{t+1} = \varepsilon'|b_t = b, \varepsilon_t = \varepsilon] \\ &= \Pr[b = b_{t+1}^{-1}(b', \varepsilon_t = \varepsilon)]\Pr[\varepsilon_{t+1} = \varepsilon'|\varepsilon_t = \varepsilon],\end{aligned}$$

where $b_{t+1}^{-1}(\cdot)$ *is the inverse function of saving* $b_{t+1}(b, \varepsilon)$*, and we have suppressed the dependence of* j *for simplicity;*

(vi) Government budget is balanced for each t

$$G_t + \mu^f \tau_t^r \delta k_t^f = \tau_t^a (p_t^a C_t^a + C_t^m) + \mu^f \tau_t^r (\pi_t^f + \pi_t^*) + \tau_t^r y_t^m + \tau_t^w (\mu^u w_t^m h_t^u + \mu^r w_t^f h_t^r),$$

(vii) Prices $\mathbf{p_t}$ clear all markets at all periods:

- *Urban Labor Market:*
$$\mu^u \int \varepsilon_t^u h_t^u d\Gamma_t^u(b^u, \varepsilon^u) = h_t^m.$$
- *Rural Labor Market:*
$$\mu^r \int h_t^r d\Gamma_t^r(b^r, \varepsilon^r) = \mu^f (h_t^a + h_t^*).$$
- *Capital Market:*
$$\mu^u \int b_{t+1}^u d\Gamma_t^u(b^u, \varepsilon^u) + \mu^r \int b_{t+1}^r d\Gamma_t^r(b^r, \varepsilon^r) = k_{t+1}^m.$$
- *Food:*
$$C_t^a = \mu^r \int z^a \varepsilon_t^r (1 - h_t^r)^{1-\alpha^a} d\Gamma_t^r(b^r, \varepsilon^r) + \mu^f z^a (h_t^a)^{1-\alpha^a}.$$
- *Services:*
$$C_t^s = \mu^u \int z^s (1 - h_t^u)^{\alpha^s} d\Gamma_t^u(b^u, \varepsilon^u).$$
- *Manufacturing Goods:*

  $$C_t^m + (k_{t+1}^m + \mu^f k_{t+1}^f) + (1 - \delta)(k_t^m + \mu^f k_t^f) + G_t = z^m (k_t^m)^{\alpha^m} (h_t^m)^{1-\alpha^m} + \mu^f R_t^*,$$

  where
  $$R_t^* = z^* (k_t^f)^{\alpha_1^*} (h_t^*)^{\alpha_2^*},$$
  is the revenue from the export sector.

A.4 *Computing the Transition Path*

Suppose that we are solving for N paths with length T, for which we use $\mathcal{P} \subseteq \prod_{i=1}^N \mathbb{R}_i^T$ to denote the space of all paths where i is the index for different prices. Then computationally, the solution trims down to the construction of two operators: an operator $\{\mathcal{T} : \mathcal{P} \to \mathcal{P}\}$ that updates the price paths, and an operator $\{\mathcal{E} : \mathcal{P} \to \prod_{i=1}^N \mathbb{R}_i^T\}$ that calculates the excess demands in all markets at all periods given the price paths. In most literature, a tatonnment algorithm based on fixed point iteration is invoked to construct the operator $\mathcal{T}$.[36] We cannot do it for our case because there is no straight forward way to update p_t^a and p_t^s. As a result, we need to rely on standard nonlinear solver to find the prices. The problem is non-trivial since for our case with 5 prices and (say) 30 periods

[36] This algorithm is initially proposed by Auerbach and Kotlikoff (1987), see Judd (1998) for a textbook treatment.

in transition, we would need to solve for a nonlinear system with $30 \times 5 = 150$ unknowns. Here we again use `hybrd` from the MINPACK.[37]

In practice, because the convergence to the new steady state is asymptotical, numerically we can only solve for an approximation of the actual transition path. The length of the transition is determined in a similar spirit as the way a boundary problem of differential equations is solved, with the distributions of households at the new steady state as the boundary condition. For a T long enough, the convergence to the second steady state is almost guaranteed as long as all the prices stay within the proximity of the second steady state.[38] Hence we say that we have found an approximation to the true transition path, when the distributions of households at $T-1$ (recall that we have assumed that in period T the economy is at the new steady state) are close to those in the new steady state $\Gamma_n^j(b^j, \varepsilon^j), j \in \{u, r\}$ by some tolerance level ε_d under certain norm:

$$\max\{\|\Gamma_{T-1}^u(b^u, \varepsilon^u) - \Gamma_n^u(b^u, \varepsilon^u)\|, \|\Gamma_{T-1}^r(b^r, \varepsilon^r) - \Gamma_n^r(b^r, \varepsilon^r)\|\} < \varepsilon_d. \tag{A.10}$$

With these in mind, we now lay out the pseudo-code that solves the transition path. We omit all the details that are the same with the steady state case.

1. Choose the transition length T.
2. Make initial guesses of the prices paths $\{\mathbf{P}_t\}_{t=1}^T = \{p_t^s, p_t^a, w_t^m, w_t^f, r_t\}_{t=1}^T$, where $\mathbf{P}_1 = \mathbf{P}_o$ and $\mathbf{P}_T = \mathbf{P}_n$.
3. Use the same algorithm in Appendix A.2 to solve for the steady state equilibrium at $t = 1$ and $t = T$. Let the corresponding stationary distributions of households be $\Gamma_1^j(\cdot) = \Gamma_o^j(\cdot)$ and $\Gamma_T^j(\cdot) = \Gamma_n^j(\cdot)$, where $j = u, r$.
4. Given $\mathbf{P}_t$, starting from $V_T^u = V_n^u$ and $V_T^r = V_n^r$, solve the households' problems (A.8) and (A.9) by backward induction to get $\{V_t^u, V_t^r\}_{t=2}^{T-1}$. In the computation, we again use FSQP and PCHI.
5. Given $\mathbf{P}_t$, starting from $k_T^f = k_n^f$, solve the large farmer's problem by backward induction using (A.4) and (A.5) to get $\{\mathbf{c}_t^f, k_t^f\}_{t=2}^{T-1}$. We use `hybrd` to solve (A.4) and (A.5).
6. Starting from $\Gamma_1^j(\cdot)$, simulate the model forward using the policy functions solved in Steps 3 and 4 until $t = T-1$. Calculate the aggregate demand and supply for the first five markets in Definition 2 in each period t. If the demands do not equal to supplies, return to Step 2 and make a new guess for $\{\mathbf{P}_t\}_{t=2}^{T-1}$. The Powell's Hybrid Method is again used here to update the prices.
7. For the prices paths $\{\mathbf{P}_t\}_{t=1}^T$ that clear all the markets, compare Γ_{T-1}^u and Γ_{T-1}^r with Γ_n^u and Γ_n^r respectively. If Equation (A.10) does not hold, increase T and repeat the whole algorithm until it is satisfied for given ε_d.

[37] Notice that in this case, the Jacobian consists of 150×150 elements and expands exponentially with T. The diagonally-dominating issue mentioned in footnote 35 still exists here. Because the prices in period t has also first order effects on the markets of periods s close to t, the problem is elevated. The algorithm hence is more sensitive to the initial guesses provided.

[38] Note that we are essentially at Step 5 of the algorithm to solve the steady state once the prices stabilize.

B. *Decomposing the Welfare Components*

In this section, we compute the consumption equivalence of welfare changes caused by tax reforms. For preferences that are homothetic, we can directly back out the consumption equivalence changes from the value function. We first show how this is done for any two levels of indirect utilities in Section B.1. We then explain how the welfare effects are decomposed into aggregate and distributional components following the idea of Domeij and Heathcote (2004) in Section B.2.

B.1 *Consumption Equivalence of Welfare Changes*

The economy begins at an initial steady state. The individual states are $x = \{b, \varepsilon\}$ and the joint distribution in the initial steady state is $\Gamma_0(b, \varepsilon)$. We use $V_0(x)$ and $V_1(x)$ for the value functions before and after the reform. Similarly, we denote the consumption and saving policy functions at time t by $c_t^{a,j}(x), c_t^{m,j}(x), c_t^{s,j}(x)$ and $b_t'^{,j}(x)$, where $j = 0, 1$ refers respectively to the scenario without and with the reform. If we let the conditional probability of x_{t+s} with respect to x_t be $\pi^j(x_{t+s}|x_t)$ for $j = 0, 1$, then the indirect utilities of a household with state x_0 in the initial steady state with and without the reform in the future are

$$V_0(x_0) = \sum_{s=0}^{\infty} \beta^s \left\{ \sum_{x_s|x_0} \pi^0(x_s|x_0)[\log c_s^{a,0}(x_s) + \gamma \log c_s^{m,0}(x_s) + \psi \log c_s^{s,0}(x_s)] \right\},$$

and

$$V_1(x_0) = \sum_{s=0}^{\infty} \beta^s \left\{ \sum_{x_s|x_0} \pi^1(x_s|x_0)[\log c_s^{a,1}(x_s) + \gamma \log c_s^{m,1}(x_s) + \psi \log c_s^{s,1}(x_s)] \right\}.$$

The consumption equivalence of welfare change λ for a household with state x_0 solves the following equation:

$$V_1(x_0) = \sum_{s=0}^{\infty} \beta^s \left\{ \sum_{x_s|x_0} \pi^0(x_s|x_0)[\log(1+\lambda) c_s^{a,0} + \gamma \log(1+\lambda) c_s^{m,0} + \psi \log(1+\lambda) c_s^{s,0}] \right\},$$

where we have suppressed the dependence on x_s for simplicity. By the property of logarithmic function and using the fact that $\sum_{x_s|x_0} \pi^0(x_s|x_0) = 1$, we have

$$V_1(x_0) = \frac{(1+\gamma+\psi)\log(1+\lambda)}{1-\beta} + V_0(x_0), \tag{B.11}$$

where we have summed the geometric sequence. From Equation (B.11), we see immediately that λ is independent of x_0. Integrating (B.11) with respect to $\Gamma_0(x)$, we have

$$1 + \lambda = \exp\left\{ (W_1 - W_0) \frac{1-\beta}{1+\gamma+\psi} \right\}, \tag{B.12}$$

where

$$W_j = \int V_j(x_0) d\Gamma_0(x_0), \quad j = 0, 1,$$

are the average welfare for a utilitarian government.

In principle, Equation (B.12) can be used to back out the consumption equivalence changes between any two scenarios. Notice that to do the conversion, only the distribution in the initial steady state and the value functions under different scenarios are needed. For non-homothetic preferences, because we do not have closed form solution for Equation (B.11) [and hence (B.12)], we would have to rely on Monte Carlo simulations to numerically find the value of λ.

B.2 *Aggregate and Distributional Components*

Decomposing λ into aggregate and distributional components essentially is determined by how we construct the indirect utility when the consumptions are scaled. The fact that there are three sectors in our model causes slight deviation from the standard one sector case. Specifically, we scale the consumption of each good by the corresponding aggregate consumption level of it to construct the indirect utility $\hat{V}_1(x_0)$.[39] If we let the aggregate consumptions of the three goods under different scenarios be $C_t^{a,j}, C_t^{m,j}$ and $C_t^{s,j}, j = 0, 1$, then the indirect utility corresponding to the aggregate component is

$$\begin{aligned}\hat{V}_1(x_0) &= \sum_{s=0}^{\infty} \beta^s \left\{ \sum_{x_s|x_0} \pi^0(x_s|x_0) \left[\log\left(\frac{C_s^{a,1}}{C_s^{a,0}}\right) c_s^{a,0} + \gamma \log\left(\frac{C_s^{m,1}}{C_s^{m,0}}\right) c_s^{m,0} + \psi \log\left(\frac{C_s^{s,1}}{C_s^{s,0}}\right) c_s^{s,0} \right] \right\} \\ &= \sum_{s=0}^{\infty} \beta^s \left[\log\left(\frac{C_s^{a,1}}{C_s^{a,0}}\right) + \gamma \log\left(\frac{C_s^{m,1}}{C_s^{m,0}}\right) + \psi \log\left(\frac{C_s^{s,1}}{C_s^{s,0}}\right) \right] + V_0(x_0),\end{aligned}$$

where again, the dependence on x_s is suppressed. Using the similar logic in deriving Equation (B.11), we can solve for the aggregate component $\hat{\lambda}$ by

$$1 + \hat{\lambda} = \exp\left\{ (\hat{W}_1 - W_0) \frac{1-\beta}{1+\gamma+\psi} \right\},$$

where

$$\hat{W}_1 = \int \hat{V}_1(x_0) d\Gamma_0(x_0).$$

The distributional component $\tilde{\lambda}$ is defined as a residual such that

$$(1 + \hat{\lambda})(1 + \tilde{\lambda}) = 1 + \lambda. \tag{B.13}$$

[39] Alternatively, we can scale consumptions on all goods up by the overall aggregate consumption. The results are very similar due to the log-linear assumption. The only significant difference is for the VAT case, since the tax directly alters relative prices.

For steady state comparisons, if we let the average utility in the initial and end steady state be W_0 and W_1, then the consumption equivalence λ_{ss} is such that

$$1+\lambda_{ss}=\exp\left\{(W_1-W_0)\frac{1-\beta}{1+\gamma+\psi}\right\}.$$

The aggregate component $\hat{\lambda}_{ss}$ is computed by assuming that the distribution is the same in the two steady states and the individual consumptions of the three goods are only adjusted respectively by the ratio of the corresponding aggregate consumptions. Hence

$$1+\hat{\lambda}_{ss}=\exp\left\{\frac{1}{1+\gamma+\psi}\left(\log\frac{C_1^a}{C_0^a}+\gamma\log\frac{C_1^m}{C_0^m}+\psi\log\frac{C_1^s}{C_0^s}\right)\right\}.$$

The distributional component is then defined residually as before:

$$(1+\tilde{\lambda}_{ss})(1+\hat{\lambda}_{ss})=1+\lambda_{ss}.$$

In practice, we do the decompositions for the urban area, rural area, and the economy as a whole respectively by using different $\Gamma_0(x)$s.

C. *The Model without Idiosyncratic Risks*

In the version of the model with no idiosyncratic risks, we set the idiosyncratic shocks $\{\varepsilon_t^r,\varepsilon_t^u\}$ to one. Due to the homothetic utility function (1) we choose, the model features aggregation [Chatterjee (1994), Azzimonti, de Francisco and Krusell (2008)]. This means that on the aggregate level, the model is indistinguishable from one with representative agents. Using this property, when solving the model without idiosyncratic risks, we first solve the representative agent version of the model to characterize the equilibrium prices paths. We then solve the dynamic programming problem given the equilibrium prices, and do welfare decompositions thereafter. Notice that one immediate corollary of aggregation is that we cannot separately solve for the consumption and savings of rural and urban households, but rather we can only pin down the average levels $C_t^H=\mu^u C_t^u+\mu^r C_t^r$ and $b_{t+1}^H=\mu^u b_{t+1}^u+\mu^r b_{t+1}^r$. With this in mind, we present the pseudo code of the algorithm that solves the prices paths of the representative agents model.

1. Choose the transition length T.
2. Set initial guesses for $\{p_t^s,p_t^a\}_{t=1}^T$, with $t=1$ and $t=T$ corresponds to the two steady states respectively.
3. Set initial guesses for $\{r_t\}_{t=1}^T$ *conditional on* the guesses of $\{p_t^s,p_t^a\}_{t=1}^T$.
4. Using the first order conditions of the manufacturing firm to back out urban wages and capital-labor ratio $\{w_t,\kappa_t^m\}_{t=1}^T$.
5. The labor supply decision of the urban household and the urban labor market clearing condition can then be used to calculate capital demand $\{k_t^m\}_{t=1}^T$.

6. Given interest rates, the Euler equations of the households (notice that here we do not distinguish between rural and urban households for reasons explained before) can be used to solve average household consumption $\{C_t^H\}_{t=2}^{T-1}$ by backward induction from C_T^H.

7. With the average consumption of households, if we can find average income $\{I_t^H\}_{t=2}^{T-1}$, using the fact that the economy is in steady state at $t = 1$ (so b_2^H is given), we can then simulate forward the sequence of capital supply $\{b_{t+1}^H\}_{t=2}^{T-1}$ using household budget constraints. The sequence of $\{b_{t+1}^H\}_{t=1}^{T-2}$ can then be used to compare with $\{k_t^m\}_{t=2}^{T-1}$ to pin down the interest rates during the transition $\{r_t\}_{t=2}^{T-1}$. Specifically, the following steps are included.

   (a) Use the urban area allocations to solve the urban income

   $$I_t^u = p_t^s z^s (1 - h_t^u)^{1-\alpha^s} + (1 - \tau_t^w) w_t^m h_t^u.$$

   (b) Now we solve for the rural allocations by backward induction. For each t, since $\{c_{t+1}^f, k_{t+1}^f, h_{t+1}^*\}$ are known, the Euler equation of the large farmer can be used to solve c_t^f. With c_t^f, the budget constraint of the large farmer defines a equation with 4 unknowns $\{k_t^f, h_t^a, h_t^*, w_t^f\}$. The labor demand functions of the large farmers provide another two equations, with the last one filled by substituting the rural labor market clearing condition to rural household's labor supply decision. Hence, at any period t, $\{k_t^f, h_t^a, h_t^*, w_t^f\}$ can be jointly solved together and the process can be extended from $t = T - 1$ to $t = 2$.

   (c) With the rural area allocation, we are able to solve the rural income

   $$I_t^r = p_t^a z^a (1 - h_t^r)^{1-\alpha^a} + (1 - \tau_t^w) w_t^f h_t^r,$$

   which in addition leads to average household income $I_t^H = \mu^u I_t^u + \mu^r I_t^r$.

   (d) We then update the guesses for $\{r_t\}_{t=2}^{T-1}$ in Step 3 until $\{b_{t+1}^H\}$ equal to $\{k_t^m\}$.

8. The average consumption $\{c_t^H\}_2^{T-1}$ can then be used to solve average consumption for food and services. We update the guesses for $\{p_t^s, p_t^a\}_{t=2}^{T-1}$ in Step 2 until the food and services goods markets are cleared.

9. With all the prices clear markets, compare the simulated b_T^H with that in the new steady state. If the difference is large, return to Step 1 and increase the transition length T.

With the equilibrium prices paths, we can use standard value function iteration technic to solve the indirect utility of households with different level of wealth. Specifically, the corresponding Bellman equations now are[40]

$$V^r(b) = \max_{\{C^r, b^{r'}, h^r\}} \left\{ u(C^r) + \beta V^r(b^{r'}) \right\}$$

$$s.t.$$

$$C^r + b^{r'} = (1 - \tau^w) w^f h^r + p^a z^a (1 - h^r)^{1-\alpha^a} + (1 + r) b^r,$$

[40] We write the steady state case for illustration, the extension of the two problems to transition is straight forward.

and

$$V^u(b) = \max_{\{C^u, b^{u'}, h^u\}} \left\{ u(C^u) + \beta V^u(b^{u'}) \right\}$$

$$s.t.$$

$$C^u + b^{u'} = (1-\tau^w) w^m h^u + p^s z^s (1-h^u)^{1-\alpha^s} + (1+r) b^u.$$

We then use the value functions to calculate the welfare decomposition following broadly the procedures in Appendix B.2 with the following caveat.

Since aggregation holds with the idiosyncratic risks shut down, the wealth distribution in this version of the model is indeterminate. For this reason, we assume that the rural and urban wealth distributions have the "same shape" as they are in the benchmark case. We say the "same shape" because the levels of aggregate capital in the equilibrium are different in the two cases. In particular, for the status quo steady state, the capital in no risk case is approximately 40% higher than the benchmark case. Hence, directly migrating the distributions from the benchmark case to the no risk case would cause inconsistency. As a result, we shift (for instance in this case by a ratio of 1.4) the wealth level in the benchmark distribution such that the aggregated capital equals that in the representative agents model.

D. *Proofs*

In this section, we prove the two results of the static model.

Result 1. *The urban-rural income gap is increasing in* τ^a.

Proof. Following our assumptions, the total income of rural households is p^a, while that of the urban households is

$$I^u = w^m(1-h^s) + p^s y^s = z^m \left[1 - \frac{(p^s)^2}{4(z^m)^2} \right] + \frac{(p^s)^2}{2z^m} \tag{D.14}$$

$$= z^m + \frac{(p^s)^2}{4z^m}.$$

Standard optimization technic thus yields the demand function for households in the rural area

$$c_r^a = \frac{1}{3(1+\tau^a)}, \qquad c_r^m = \frac{p^a}{3(1+\tau^m)}, \qquad c_r^s = \frac{p^a}{3p^s},$$

and in the urban area

$$c_u^a = \frac{I^u}{3(1+\tau^a)p^a}, \qquad c_r^m = \frac{I^u}{3(1+\tau^m)}, \qquad c_r^s = \frac{I^u}{3p^s}.$$

Given the solution to the households' problems, the food market clearing condition is

$$\underbrace{\frac{2}{3} \cdot \frac{1}{3(1+\tau^a)}}_{\text{Rural Demand}} + \underbrace{\frac{1}{3} \cdot \frac{I^u}{3(1+\tau^a)p^a}}_{\text{Urban Demand}} = \frac{2}{3}.$$

Rearranging terms with some algebra yields

$$I^u = 2[3(1+\tau^a)-1]p^a, \tag{D.15}$$

where it can be immediately seen that the urban-rural income gap I^u/p^a is increasing in the value added tax on food τ^a. □

Result 2. *If the government uses the tax revenue collected through value added tax to purchase the same good, then value added tax has zero efficiency cost.*

Proof. We use the value added tax on the manufacturing good τ^m as an example. The same logic goes through if instead we assume $\tau^a \neq 0$ and $\tau^m = 0$, but all tax revenue is used by the government to purchase agricultural good. The output in the rural area is always 2/3 unit of agricultural good, and will not be distorted by any taxes. Hence the key here is to show that when $\tau^m \neq 0$, production in the urban area is the same as in the first best.

For each urban household, the labor supply decision equalizes the marginal returns from the formal and informal sector:

$$(1-\tau^w)w_m = \frac{\partial p^s y^s}{\partial h^s} = \frac{1}{2}p^s(h^s)^{-1/2},$$

which further yields

$$h^s = \frac{(p^s)^2}{4[(1-\tau^w)z^m]^2}. \tag{D.16}$$

According to Equation (D.16), it is equivalent to showing that p^s does not change with τ^m, which we now prove.

Combining Equations (D.14) and (D.15), we have

$$\frac{p^a}{p^s} = \frac{1}{M}\left[\frac{z^m}{p^s} + \frac{p^s}{4z^m}\right], \tag{D.17}$$

where we let $M = 2[3(1+\tau^a)-1]$ to simplify notation. Like before, after substituting the households decisions, the services market clearing condition becomes

$$\underbrace{\frac{2}{3}\cdot\frac{p^a}{3p^s}}_{\text{Rural Demand}} + \underbrace{\frac{1}{3}\cdot\frac{I^u}{3p^s}}_{\text{Urban Demand}} = \frac{1}{3}\cdot\frac{p^s}{2z^m}. \tag{D.18}$$

Further substituting Equations (D.14) and (D.17) into (D.18), after rearranging terms and some algebra, we arrive at the expression of services good

$$p^s = \left[\frac{4(M+2)}{5M-2}\right]^{\frac{1}{2}} z^m, \tag{D.19}$$

which does not depend on τ^m. □

TABLE E.1—THE WELFARE COSTS WHEN RURAL HOUSEHOLDS ARE VALUED MORE

	VAT		CIT		PIT	
	Equal	Rural	Equal	Rural	Equal	Rural
Total	−4.01%	−4.51%	−1.94%	−1.90%	−3.21%	−3.19%
Aggregate	−2.58%	−3.34%	−1.91%	−1.85%	−3.37%	−3.20%
Distributional	−1.46%	−1.21%	−0.02%	−0.05%	0.17%	0.01%

E. *The Government Objective Function*

In the benchmark experiments, we assume that the government is utilitarian. Table E.1 presents the results when the weights the government assigns to each rural household are doubled. We find that the changes are small because Ethiopia already features a large rural population (69%).

F. *A Toolkit to Implement the Quantitative Exercises*

To bridge the academic findings with policy prescriptions, a toolkit with graphical interfaces is developed in companion with the paper. Figure A1 provides a screenshot of the main interfaces. The toolkit allows users with limited knowledge of programming language to calibrate the model to different target economies, and use the calibrated model to conduct analyses similar to those in the main text. The toolkit also exports a large number of intermediate results in Microsoft Excel `.xlsx` format, which the users can exploit to produce analyses do not feature into the paper. The latest version of the toolkit (currently version 2.0) is distributed through GitHub at https://github.com/zjutangxin/toolkit_ver2_dist. Notice that reuse of the toolkit does not imply any endorsement of the research and/or product. Any research presented should not be reported as representing the views of the IMF, its Executive Board, member governments, and the Department for International Development (DFID) of the United Kingdom.

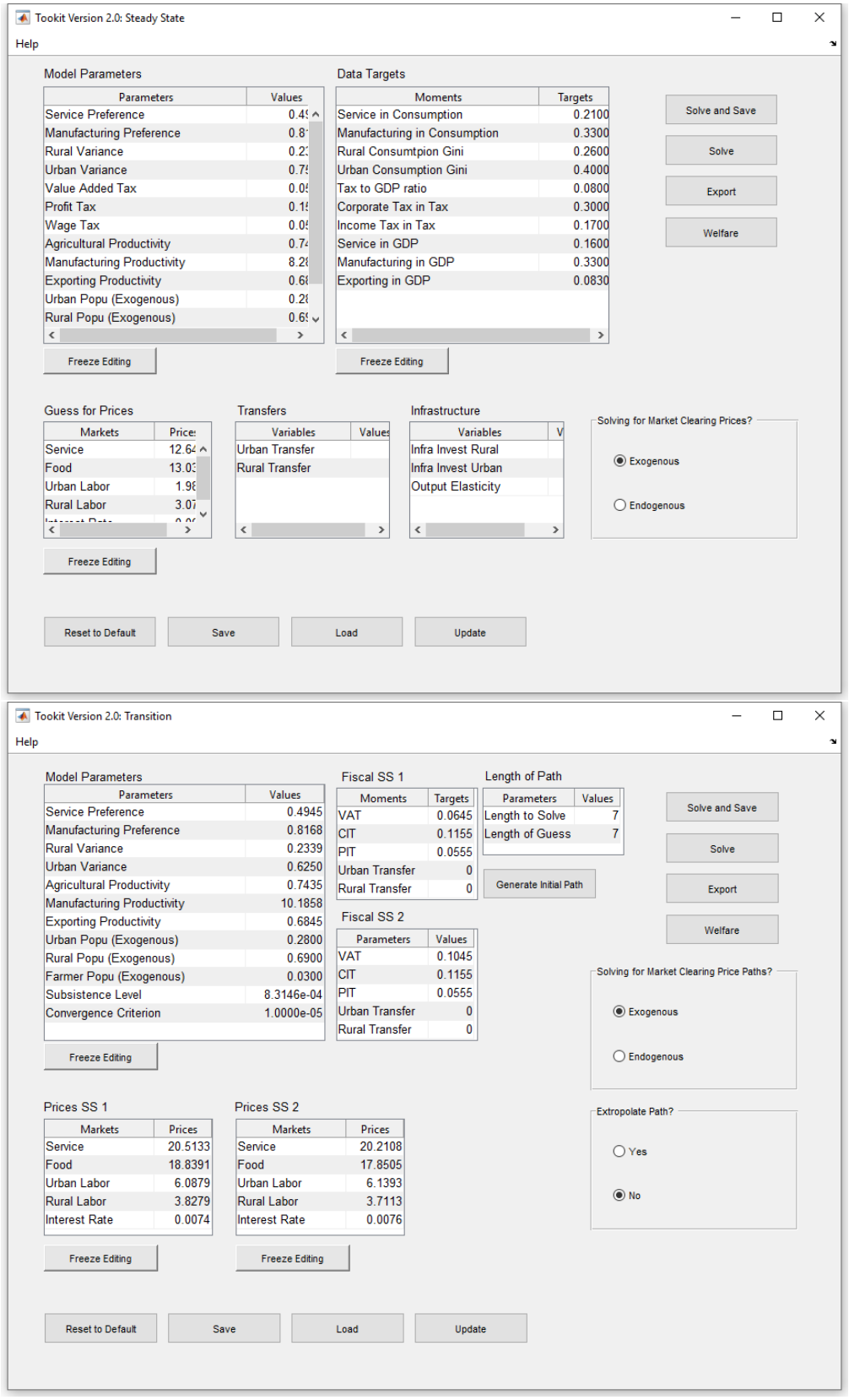

FIGURE A1. STANDARD USER INTERFACES OF THE TOOLKIT

Upper Panel: Steady State; Lower Panel: Transitional Dynamics.

REFERENCES

Adamopoulos, Tasso, and Diego Restuccia. 2014. "The Size Distribution of Farms and International Productivity Difference." *American Economic Review*, 104(6): 1667–1697. 9, 10, 15

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *The Quarterly Journal of Economics*, 109(3): 659–684. 4, 7, 31

Aiyagari, S. Rao. 1995. "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting." *Journal of Political Economy*, 103(6): 1158–1175. 6

Allen, Treb, and David Atkin. 2017. "Volatility and the Gains from Trade." *Manuscript*. 7, 10

Anagnostopoulos, Alexis, and Qian Li. 2013. "Consumption Taxes and Precautionary Savings." *Economics Letters*, 119(3): 238–242. 19

Anagnostopoulos, Alexis, Eva Cárceles-Poveda, and Danmo Lin. 2012. "Dividend and Capital Gains Taxation under Incomplete Markets." *Journal of Monetary Economics*, 59(7): 599–611. 5

Anderson, Kym, Gordon Rausser, and Johan Swinnen. 2013. "Political Economy of Public Policies: Insights from Distortions to Agricultural and Food Markets." *Journal of Economic Literature*, 51(2): 423–477. 9

Angeletos, George-Marios. 2007. "Uninsured Idiosyncratic Investment Risk and Aggregate Saving." *Review of Economic Dynamics*, 10(1): 1–30. 7

Auerbach, Alan J., and Laurence J. Kotlikoff. 1987. *Dynamic Fiscal Policy.* Cambridge:Cambridge University Press. 35

Azzimonti, Marina, and Pierre Yared. 2019 forthcoming. "The Optimal Public and Private Provision of Safe Assets." *Journal of Monetary Economics*. 8

Azzimonti, Marina, Eva de Francisco, and Per Krusell. 2008. "Aggregation and Aggregation." *Journal of the European Economic Association Papers and Proceedings*, 6(2-3): 381–394. 25, 39

Bakış, Ozan, Barış Kaymak, and Markus Poschke. 2015. "Transitional Dynamics and the Optimal Progressivity of Income Redistribution." *Review of Economic Dynamics*, 18(3): 679–693. 5

Benhabib, Jess, Richard Rogerson, and Randall Wright. 1991. "Homework in Macroeconomics: Household Production and Aggregate Fluctuations." *Journal of Political Economy*, 99(6): 1166–1187. 6

Burgess, Robin, Olivier Deschenes, Dave Donaldson, and Michael Greenstone. 2017. "Weather, Climate Change, and Death in India." *Manuscript*. 10

Carroll, Christopher D. 1997. "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis." *The Quarterly Journal of Economics*, 112(1): 1–55. 31

Carroll, Christopher D., and Miles S. Kimball. 1996. "On the Concavity of Consumption Function." *Econometrica*, 64(4): 981–992. 27

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica*, 54(3): 607–622. 7

Chari, V. V., and Patrick J. Kehoe. 1999. "Optimal Fiscal and Monetary Policy." In *Handbook of Macroeconomics*. Vol. 1C, , ed. John B. Taylor and Michael Woodford, Chapter 26, 1671–1745. North Holland:Elsevier B.V. 7

Chatterjee, Satyajit. 1994. "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model." *Journal of Public Economics*, 54(1): 97–119. 25, 39

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis. 2017. "The Effects of Land Markets on Resource Allocation and Agricultural Productivity." *NBER Working Paper No. 24034*. 10

Chetty, Raj, and Adam Szeidl. 2007. "Consumption Commitments and Risk Preferences." *The Quarterly Journal of Economics*, 122(2): 831–877. 9

Coleman, Wilbur John II. 2000. "Welfare and Optimum Dynamic Taxation of Consumption and Income." *Journal of Public Economics*, 76(1): 1–39. 7

Conesa, Juan Carlos, Bo Li, and Qian Li. 2018. "Welfare Implications of Switching to Consumption Taxation." *Manuscript*. 5

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing Capital? Not a Bad Idea After All!" *American Economic Review*, 99(1): 25–48. 6

Dávila, Julio, Jay H. Hong, Per Krusell, and José-Víctor Ríos-Rull. 2012. "Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks." *Econometrica*, 80(6): 2431–2467. 22

Domeij, David, and Jonathan Heathcote. 2004. "On the Distributional Effects of Reducing Capital Taxes." *International Economic Review*, 45(2): 523–554. 5, 6, 16, 17, 18, 19, 33, 37

Donovan, Kevin. 2018. "Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences." *Manuscript*. 5, 7, 9, 10, 24, 25

Duflo, Esther, and Rohini Pande. 2007. "Dams." *The Quarterly Journal of Economics*, 122(2): 601–646. 10, 11

Eastwood, Robert, Michael Lipton, and Andrew Newell. 2010. "Farm Size." In *Handbook of Agricultural Economics*. Vol. 4, , ed. Prabhu L. Pingali and Robert E. Evenson, Chapter 65, 3323–3397. Burlington:Elsevier BV. 8, 11

Feng, Ying, David Lagakos, and James E. Rauch. 2018. "Unemployment and Development." *NBER Working Paper No. 25171*. 29

Flodén, Martin. 2006. "Labour Supply and Saving under Uncertainty." *Economic Journal*, 116(513): 721–737. 6

Fritsch, F. N., and J. Butland. 1984. "A Method for Constructing Local Monotone Piecewise Cubic Interpolants." *SIAM Journal on Scientific and Statistical Computing*, 5(2): 300–304. 32

Fritsch, F. N., and R. E. Carlson. 1980. "Monotone Piecewise Cubic Interpolation." *SIAM Journal on Numerical Analysis*, 17(2): 238–246. 32

Gollin, Douglas, David Lagakos, and Michael E. Waugh. 2014. "The Agricultural Productivity Gap." *Quarterly Journal of Economics*, 129(2): 939–993. 4

Gollin, Douglas, Stephen L. Parente, and Richard Rogerson. 2004. "Farm Work, Home Work and International Productivity Differences." *Review of Economic Dynamics*, 7(4): 827–850. 6

Gollin, Douglas, Stephen L. Parente, and Richard Rogerson. 2007. "The Food Problem and the Evolution of International Income Levels." *Journal of Monetary Economics*, 54(4): 1230–1255. 8

Heathcote, Jonathan, Kjetil Storesletten, and Gianluca Violante. 2014. "Consumption and Labor Supply with Partial Insurance: An Analytical Framework." *American Economic Review*, 104(7): 2075–2126. 7

Heathcote, Jonathan, Kjetil Storesletten, and Gianluca Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework." *The Quarterly Journal of Economics*, 132(4): 1693–1754. 7

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2008. "Insurance and Opportunities: A Welfare Analysis of Labor Market Risk." *Journal of Monetary Economics*, 55(3): 501–525. 7

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi. 2014. "Growth and Structural Transformation." In *Handbook of Economic Growth*. Vol. 2B, , ed. Philippe Aghion and Steven N. Durlauf, Chapter 6, 854–941. North Holland:Elsevier B.V. 4

Judd, Kenneth L. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, 28(1): 58–83. 7

Judd, Kenneth L. 1998. *Numerical Methods in Economics*. Cambridge, Massachusetts and London, England:MIT Press. 35

Kaplan, Greg, and Giovanni L. Violante. 2010. "How Much Consumption Insurance beyond Self-

Insurance?" *American Economic Journal: Macroeconomics*, 2(4): 53–87. 7

Lagakos, David, and Michael E. Waugh. 2013. "Selection, Agriculture, and Cross-Country Productivity Differences." *American Economic Review*, 103(2): 948–980. 4, 9

Lagakos, David, Mushfiq Mobarak, and Michael E. Waugh. 2017. "The Welfare Effects of Encouraging Rural-Urban Migration." *Manuscript*. 8, 10, 24, 29

La Porta, Rafael, and Andrei Shleifer. 2014. "Informality and Development." *Journal of Economic Perspectives*, 28(3): 109–126. 4, 11

Lucas, Robert E. 1987. *Models of Business Cycles.* Cambridge, Massachussets:Basil Blackwell. 19

Lucas, Robert E.B. 1997. "Internal Migration in Developing Countries." In *Handbook of Population and Family Economics*. Vol. 1B, , ed. Mark R. Rosenzweig and Oded Stark, Chapter 13, 721–798. Elsevier Science B.V. 8

Marcet, Albert, Francesc Obiols-Homs, and Philippe Weil. 2007. "Incomplete Markets, Labor Supply and Capital Accumulation." *Journal of Monetary Economics*, 54(8): 2621–2635. 6, 25

Meghir, Costas, and Luigi Pistaferri. 2011. "Earnings, Consumption and Life Cycle Choices." In *Handbook of Labor Economics*. Vol. 4B, , ed. Orley Ashenfelter and David Card, Chapter 9. Elsevier B.V. 16

Mendoza, Enrique G., Vincenzo Quadrini, and José-Víctor Ríos-Rull. 2009. "Financial Integration, Financial Development, and Global Imbalances." *Journal of Political Economy*, 117(3): 371–416. 7

Morten, Melanie. 2019. "Temporary Migration and Endogenous Risk Sharing in Village India." *Journal of Political Economy*, 127(1): 1–46. 5, 24

Parente, Stephen L., Richard Rogerson, and Randall Wright. 2000. "Homework in Development Economics: Household Production and the Wealth of Nations." *Journal of Political Economy*, 108(4): 680–687. 6

Peralta-Alva, Adrian, Xuan Song Tam, Xin Tang, and Marina Mendes Tavares. 2018. "The Macroeconomic and Distributional Implications of Fiscal Consolidations in Low-income Countries." *IMF Working Paper (18/146)*. 9

Pijoan-Mas, Josep. 2006. "Precautionary Savings or Working Longer Hours?" *Review of Economic Dynamics*, 9(2): 326–352. 6, 25

Pingali, Prabhu L. 2007. "Agricultural Mechanization: Adoption Patterns and Economic Impact." In *Handbook in Agricultural Economics*. Vol. 3, , ed. Robert E. Evenson and Prabhu L. Pingali, Chapter 54, 2779–2805. North Holland:Elsevier B.V. 11

Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *The Economic Journal*, 37(145): 47–61. 18

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu. 2008. "Agricultural and Aggregate Productivity: A Quantitative Cross-Country Analysis." *Journal of Monetary Economics*, 55(2): 234–250. 9

Ríos-Rull, José-Víctor. 1999. "Computation of Equilibria in Heterogeneous-agent Models." In *Computational Methods for the Study of Dynamic Economies*. , ed. Ramon Marimon and Andrew Scott, 238–273. New York:Oxford University Press. 33

Roy, Andrew Donald. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers*, 3(2): 135–146. 4

Schneider, Friedrich, and Dominik H. Enste. 2000. "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature*, 38(1): 77–114. 4

Schultz, Theodore William. 1953. *The Economic Organization of Agriculture.* New York:McGraw-Hill. 9

Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron. 2001. "The Welfare Cost of Business Cycles Revisited: Finite Lives and Cyclical Variation in Idiosyncratic Risk." *European Economic Review*, 45(7): 1311–1339. 19

Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron. 2004. "Cyclical Dynamics in Idiosyncratic Labor Market Risk." *Journal of Political Economy*, 112(3): 695–717. 16

Tauchen, George. 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economic Letters*, 20(2): 177–181. 16, 31

Tombe, Trevor. 2015. "The Missing Food Problem: Trade, Agriculture, and International Productivity Differences." *American Economic Journal: Macroeconomics*, 7(3): 226–258. 8

Townsend, Robert M. 1994. "Risk and Insurance in Village India." *Econometrica*, 62(3): 539–591. 4, 5

Udry, Christopher. 1994. "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria." *The Review of Economic Studies*, 61(3): 495–526. 4, 5

Werning, Iván, and Ludwig Straub. 2018. "Positive Long Run Capital Taxation: Chamley-Judd Revisited." *Manuscript*. 7

Wu, Yiyun, Xican Xi, Xin Tang, Deming Luo, Baojing Gu, Shu Kee Lam, Peter Vitousek, and Deli Chen. 2018. "Policy Distortions, Farm Size, and the Overuse of Agricultural Chemicals in China." *Proceedings of the National Academy of Sciences of the United States of America*, 115(27): 7010–7015. 10

Young, Alwyn. 2013. "Inequality, the Urban-Rural Gap, and Migration." *The Quarterly Journal of Economics*, 128(4): 1727–1785. 4

Zhou, Jian L., André L. Tits, and Craig T. Lawrence. 1997. "User's Guide for FFSQP Version 3.7: A FORTRAN Code for Solving Constrained Nonlinear (Minimax) Optimization Problems, Generating Iterates Satisfying All Inequality and Linear Constraints." 32