We quantitatively investigate the macroeconomic and distributional impacts of fiscal consolidations in low-income countries (LICs) 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 rural-urban distinction to capture salient features of LICs. We find that overall, VAT has the least efficiency costs but is highly regressive, while PIT impacts the economy in the opposite way with CIT staying in between. Cash transfers targeting rural households mitigate the negative distributional impacts of VAT most effectively, while public investment leads to little redistribution. (JEL D31, E62, H23)

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 that are 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. A large literature has shown that taxation has implications on both efficiency and equity. One outstanding question is that in low-income countries, what is the most desirable tax instrument for fiscal consolidation that balances the efficiency and equity concerns.

In this paper, we study quantitatively the macroeconomic and distributional impacts of different tax instruments for low-income countries. The economic structure of low-income countries differs

1 Disclaimer: The research results and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the International Monetary Fund and the Department for International Development (DFID) of the United Kingdom.

Peralta-Alva: International Monetary Fund, 700 19th St NW, Washington D.C., United States, 20431 (e-mail: aperalta-alva@imf.org); Tavares: International Monetary Fund, (e-mail: marinamendestavares@gmail.com); Tam: City University of Hong Kong, Kowloon Tong, Kowloon, Hong Kong, (e-mail: xuanstam@cityu.edu.hk); Tang: International Monetary Fund and Wuhan University, (e-mail:zjutangxin@gmail.com). For helpful comments we thank Andy Berg, Yongsung Chang, Rupa Duttagupta, Hui He, Catherine Pattillo, Chris Papageorgiou, and seminar participants at the 2015 Macroeconomic Policy and Income Inequality Workshop at the IMF, Federal Reserve Bank of St. Louis, and Northeast Universities Development Consortium Conference (NEUDC) at Brown University. The authors are grateful to DFID for generous financial supports.

1 For instance, over the 1990s, Peru increased its revenue ratio from 6% to 13% of GDP, which further increased to 17% by 2010; Tanzania and Vietnam have achieved sustained increases of 4% to 5% of GDP over a period of 5 to 6 years; and Mauritania increased the tax revenue ratio by 6% of GDP between 2011 and 2014.
significantly from that of advanced economies. In particular, low-income countries usually have a large unproductive agricultural sector, face substantial informal economic activity, and observe a sharp rural-urban distinction. The unique economic structure of low-income countries implies that they need to deal with trade-offs that are different from those typically faced by developed countries. Take value added tax (consumption tax) for example. While according to classical Ramsey taxation theory, a uniform consumption tax across all goods is equivalent to non-distorting lump-sum tax [Chari and Kehoe (1999) and Coleman (2000)], in a low-income country with serious informal and shadow economy, an increase in consumption tax may worsen the informality problem. Therefore, it is important to evaluate quantitatively the consequences of various tax instruments in a model economy that captures salient features of low-income countries.

To guide our analysis, we extend the workhorse heterogeneous agents incomplete markets model [Aiyagari (1994)] to include features that are important to the study of low-income countries. Specifically, we consider a model with four sectors of different productivity—food, manufacturing, services, exporting cash crops—two regions with segmented labor markets—rural and urban—and a unified capital market. Each region has a formal and an informal labor market. For the rural area, the formal and informal labor markets are respectively cash crops and food, with those for the urban area being manufacturing and services. Each region is populated by a continuum of households who consume food, manufacturing goods, and services with food consumption subject to a subsistence requirement. Each household also faces persistent idiosyncratic productivity shocks on both markets that can only be partially insured against through 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. The government has access to three Ramsey taxes in order to raise additional revenues: value added tax (VAT) on food and manufacturing goods consumption, personal income tax (PIT) on formal wage income, and corporate income tax (CIT) on manufacturing and cash crops production.

To discipline our quantitative analysis, we calibrate our benchmark economy to Ethiopia, which we choose as representative of low-income countries. Our model can replicate well the sectoral consumption and production shares, tax structure, and regional consumption inequality observed in Ethiopia. We then use the calibrated model to evaluate the macroeconomic and distributional impacts of different fiscal consolidation schemes. Specifically, we conduct steady state comparison across revenue mobilization schemes where an additional tax revenues equal to 2% GDP in the benchmark economy are raised by VAT, PIT, and CIT respectively.

Our quantitative results show that across the three taxes, VAT leads to the least output and consumption losses of respectively 1.8% and 4% due to its non-distorting feature, but additional elements in our model also implies that some other side effects may present. It causes the expansion of the informal sector due to decrease in the relative price of informal services, and is particularly regressive in low-income countries. This latter effect comes from two channels that are not presented in standard models. First the multi-sector structure of our model implies that asymmetric changes in the prices of consumption goods would affect households of different income sources

---

2These features of low-income countries are well documented in the literature. See for instance, 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.
differentially. In particular, discouraged food demand leads to price drop that hurts rural farmers relying on food sales, thus widening the rural-urban gap. Second, the Stone-Geary preference in our model suggests that higher subsistence expenditure requirement reduces the welfare of households with lower disposable income more [Chetty and Szeidl (2007)]. On the other hand, though the efficiency costs of PIT and CIT are doubled compared to those of VAT, they have considerably better distributional impacts, as most of the tax burdens fall on rich urban households.

Inspired by these results, we then continue to investigate whether government cash transfer programs and public investment that boosts overall TFP could mitigate the negative impacts of revenue mobilization on the economy. In particular, we consider two main fiscal expenditure schemes under the VAT increase case, where 50% of the additional tax revenues are used respectively to finance a cash transfer program that targets all rural households, and investment in infrastructure equal to a 1.74% increase in TFP. Our results show that targeted cash transfer program is able to make up for 40% of the total reduction in consumption, and can completely overturn the negative distributional impacts of VAT. Infrastructure investment, on the other hand, has virtually no additional impacts on the distribution of economic resources. In addition, we show that cash transfer programs are most effective when targeted at rural households as opposed to a universal basic income across all households. This further highlights the importance of treating rural and urban areas distinctly in a uniform framework.

Related Literature.—This paper is most closely related to the extensive literature studying the macroeconomic and distributional impacts of different tax instruments quantitatively in incomplete markets models. Early works in this literature include Ventura (1999), Altig et al. (2001), Domeij and Heathcote (2004), and Nishiyama and Smetters (2005), which all focus on the United States. While later work extends the literature to the discussion of progressivity and normative property [Conesa and Krueger (2006), Conesa, Kitao and Krueger (2009), Heathcote, Storesletten and Violante (2017a), Heathcote, Storesletten and Violante (2017b)], dividend and capital income taxes [Anagnostopoulos, Cárcceles-Poveda and Lin (2012)], and Mirrleesian tax [Heathcote and Tsujiyama (2017)], our paper maintains the Ramsey tax assumption but extend the analysis to the environment of low-income countries. We show that the multi-sector multi-region structure of low-income countries could significantly change the distributional impacts of taxes comparing to their effects in advanced economies like the United States. In this regard, we view our paper as most closely related to Domeij and Heathcote (2004). In a broad sense, our paper is also related to empirical work studying the revenue mobilization issues in low-income countries [Keen and Lockwood (2010) and Ufier (2014)], though we take a different quantitative approach.

The remainder of the paper proceeds as follows. We first describe the model in Section I and then calibrate its benchmark version in Section II. Section III contains the main results, and we conclude in Section IV. Several relevant issues such as the computational algorithm, are included in the Appendix.

---

3There is also a companion literature that is built under complete markets. See for instance, Chamley (1986), Judd (1985), Correia (2010), and more recently Werning and Straub (2015).
I. 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 the Federal Democratic Republic of 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. Production in Ethiopia is overwhelmingly of a subsistence nature, and a large part of commodity exports are provided by the small agricultural cash crop sector, with coffee as the largest foreign exchange earner, and its flower industry becoming a new source of revenue in recent years. Ethiopia is heavily reliant on imported oil, and the country lacks sufficient foreign exchange.

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. Each region is populated by a continuum of infinitely lived households with the population shares denoted by $\mu_r$ and $\mu_u$. Each household is endowed with one unit of time, which is divided between the regional formal labor markets and informal self-employment. We assume that labor markets in the two regions are segmented, and households cannot migrate.

In the rural area, households work informally on their own arable land to grow food. A share $\mu_f$ of large farmers hire labor in the formal market to produce both food and cash crops. In the urban area, households work informally to provide services goods and a representative neoclassical firm hires labor in the formal market to produce manufacturing goods. Households’ labor use of all purposes are subject to idiosyncratic productivity shocks. 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, and services goods optimally. We maintain the standard assumption in the literature that savings are turned into capital of equivalent value. 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 gross 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. The manufactured goods are used for consumption and in manufacturing goods and cash crops production. This broadly captures 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]

---

4We 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 [Kennan and Walker (2011) and Lagakos, Mobarak and Waugh (2017)].
Tombe (2015).]

The government has access to three Ramsey (linear) taxes: valued added tax (VAT, $\tau_a$) on food and manufacturing goods, personal income tax (PIT, $\tau^w$) on households’ income from formal markets, and corporate income tax (CIT, $\tau^r$) 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 can spend the tax revenues on non-productive public expenditure $G$, transfers to urban and rural households $T^u$ and $T^r$, and infrastructure investment $K^g$ that boosts private sector productivity. We assume that the government runs balanced budget in every period.\footnote{Adding government debt would crowd out private investment. The model would have a similar response on output, but the implications on interest rate would be different. See Appendix D for more details.}

We next describe the model more formally.

B. Preference

We assume that both types of households and the large farmer share the same preference over sequences of consumption on food, manufacturing, and services goods $c_t = (c^a_t, c^m_t, c^s_t)$:

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

where $\beta$ is the time preference. To capture the fact that poor people spend a disproportionately higher fraction of their total income on food consumption, we follow the structural transformation literature by assuming a Stone-Geary preference\footnote{See for example, Lagakos and Waugh (2013), Herrendorf, Rogerson and Valentinyi (2013), and Adamopoulos and Restuccia (2014).}:

$$u(c_t) = \log(c^a_t - \bar{\pi}) + \gamma \log c^m_t + \psi \log c^s_t,$$

where $\bar{\pi}$ is the subsistence level, $\gamma$ and $\psi$ are relative preferences over manufacturing and service goods, and subscript $t$ refers to time periods.

One implication of the Stone-Geary preference (1) is that households always purchase at least $\bar{\pi}$ units food before distribute their consumption expenditure $c^f$ over other goods. Therefore as households become richer, the share of food expenditure in $\bar{C}$ declines steadily. In addition, the remaining $\bar{C} - (1 + \tau^a)p^a\bar{\pi}$ is distributed to $c^a, c^m$ and $c^s$ by shares $1/(1 + \gamma + \psi), \gamma/(1 + \gamma + \psi)$ and $\psi/(1 + \gamma + \psi)$ respectively. 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:

$$y^a_t = z^0\varepsilon^r_t (d^r)^{\alpha^a} (1 - h^r_t)^{1-\alpha^a},$$
where $z^a$ is economy-wide agricultural productivity, $d^r$ is exogenously given land endowment of each household, $1 - h^r_t$ is the share of labor supplied informally, and $\alpha^a$ is the factor share. Similarly, by working informally, urban households produce service goods according to the production function:

$$y^s_t = z^s(1 - h^u_t)^{1 - \alpha^s},$$

where $z^s$ is economy-wide service productivity, $1 - h^u_t$ is the share of labor supplied informally, and $\alpha^s$ is the factor share. Notice that in the two production functions, there are two idiosyncratic shocks $\varepsilon^r_t$ and $\varepsilon^u_t$. In the context of development economics, we use idiosyncratic shock $\varepsilon^r_t$ to capture the fact that agricultural production is risky, especially in developing countries which usually lack irrigation system or hydraulic infrastructure to smooth weather shocks. It also captures variations in crop yields due to difference in individual farmers’ ability or crop choice. Likewise, $\varepsilon^u_t$ quantifies variations in demand shocks and productivity at the individual level in a parsimonious way.

**Formal Markets.**—We assume that labor productivity in formal markets for both types of households are also subject to conceptually similar idiosyncratic shocks $\tilde{\varepsilon}^r_t$ and $\tilde{\varepsilon}^u_t$. These shocks reflect the efficiency of the labor hours supplied to formal markets by households. Thus the actual pre-tax wages per unit labor are respectively $w^f_t \tilde{\varepsilon}^r_t$ and $w^m_t \tilde{\varepsilon}^u_t$ for rural and urban households. Manufacturing firms produce by the production function:

$$y^m_t = z^m(h^m_t)^{1 - \alpha^m},$$

where $z^m$ is manufacturing productivity, $h^m_t$ is the total labor units hired. Meanwhile, large farmers produce food following the same production function as (2):

$$y^{a,f}_t = z^a(h^a_t)^{1 - \alpha^a}.$$ 

We assume that the production of cash crops is modernized. This means that it requires modern farm machinery $k^f_t$ and is operated at larger scale measured in land $d^r$:

$$y^*_t = z^*(d^r)^{1 - \alpha^*}1, \quad \sum_{t \in {u,r,f}} \alpha^{*}_{1} + \alpha^{*}_{2} = 1,$$

where $z^*$ is exporting sector productivity, $d^* > d^r$ is exogenously given land, and $\alpha^*_1, \alpha^*_2$ are factor shares. Likewise, here $h^a_t$ and $h^*_t$ are also measured in efficient units. Notice that we assume that there are no risks associated with production in the formal market.

### D. Optimization Problems

To simplify the notation, we let

$$C^j_t = (1 + \tau^a)(p^a_t c^a_j + \epsilon^{m,j}_t) + p^s_t c^{s,j}_t, \quad j \in \{u,r,f\}$$

denote the total consumption expenditure for urban, rural households, and large farmers.\(^8\) Throughout it is understood that the $C^j_t$’s are always spent according to the optimal consumption bundle.

\(^7\)See for example, Allen and Atkin (2017), Burgess et al. (2017), Donovan (2017), and Lagakos, Mobarak and Waugh (2017).

\(^8\)Though we assume that large farmers also consume all kinds of goods, but 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.
Households.—Let \( b^j \) represents savings and \( \varepsilon^j = (\varepsilon^j, \tilde{\varepsilon}^j), j = u, r \). We assume that the \( \varepsilon^j \)s are Markovian processes. The recursive problem of the urban households is:

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

\[\text{s.t.}\]

\[C^u + b^u = (1 - \tau^u)\varepsilon^u w^m h^u + p^* z^* \varepsilon^u (1 - h^u)^{1-\alpha^*} + (1 + r) b^u.\]

That of the rural households is

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

\[\text{s.t.}\]

\[C^r + b^r = (1 - \tau^r) w^f \varepsilon^r h^r + p^a z^a \varepsilon^r (d^r)^{\alpha^a} (1 - h^r)^{1-\alpha^a} + (1 + r) b^r.\]

It is natural to expect that for each individual household, labor productivities on the formal and informal markets are correlated. To keep the problem computationally tractable, we assume that \( \tilde{\varepsilon}^j \) can be constructed from \( \varepsilon^j \), reducing the state variables to only \( b \) and \( \varepsilon \). To discipline our analysis, we require our model to be consistent with the empirical evidence which suggests that informal sectors are usually much less productive than formal sector [La Porta and Shleifer (2014)], and that hired labor is also much less productive in agricultural sector [Eastwood, Lipton and Newell (2010)]\(^9\). In particular, in the calibration, we construct \( \tilde{\varepsilon}^j \) from \( \varepsilon^j \) such that households with low productivity choose to work in the formal market in the rural area, but in the informal market in the urban area.\(^10\)

Notice that the two dynamic programming problems share very similar structure. With a slight abuse of notations, the solutions to these two problems will be policy functions on consumption \( c(b, \varepsilon) \), saving \( b'(b, \varepsilon) \), and labor supply to formal markets \( h(b, \varepsilon) \). The joint cumulative distribution functions of households in the rural and urban areas at the steady state are denoted respectively by \( \Gamma^r(b^r, \varepsilon^r) \) and \( \Gamma^u(b^u, \varepsilon^u) \).

Farmers and Firms.—Since we are focusing on the steady states, and we have assumed that the decisions of farms and large farmers are deterministic, their sequential optimization problems can be solved analytically given prices. Specifically, the sequential problem of the large farmer is

\[
\max \left\{ \sum_{t=0}^{\infty} \beta^t u(C_t^f) \right\}
\]

\[\text{s.t.}\]

\[C_t^f + h_{t+1}^f = (1 - \tau^r)(\pi^f_t + \pi^*_t) + (1 - \delta) h_t^f + \tau^f \delta h_t^f,\]

\[\pi^f_t = p^a z^a (d^r)^{\alpha^a} (h_t^a)^{1-\alpha^a} - w^f h_t^a,\]

\[\pi^*_t = p^* z^* (d^r)^{\alpha^*_a} (h_t^a)^{1-\alpha^*_a} - w^f h_t^a.\]

\(^9\)Empirical studies have documented various causes of such systematic difference in productivity. Human capital is usually considered as an important factor behind the formal productivity premium, while the low productivity of hired labor in agriculture is mostly considered as a consequence of moral hazard.

\(^{10}\)See Appendix A for further details.
where CIT is collected over farmers’ profits. The manufacturing firm’s problem is

\[
\max_{\{k^m_t, h^m_t\}} \left\{ \left(1 - \tau^r\right) z^m(k^m)^\alpha^m(h^m)^{1-\alpha^m} - w^m h^m - (r + \delta) k^m \right\}.
\]

Notice that CIT is imposed on firm’s revenue, and depreciation \(\delta\) is also paid by the firm.

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

\[
C^x_t = \mu^u \int \varepsilon^x_{t} u d\Gamma^u(b^u_t, \varepsilon^u_t) + \mu^r \int \varepsilon^x_{t} r d\Gamma^r(b^r_t, \varepsilon^r_t) + \mu^f \varepsilon^x_{t} f, \quad x = \{a, m, s\},
\]

the aggregate consumption of each goods,

\[
h^u = \int \tilde{\varepsilon}^u_{t} h^u_t d\Gamma^u(b^u_t, \varepsilon^u_t), \quad h^r = \int \tilde{\varepsilon}^r_{t} h^r_t d\Gamma^r(b^r_t, \varepsilon^r_t),
\]

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

\[
y^m = z^m(k^m)^\alpha^m(h^m)^{1-\alpha^m},
\]

the total revenue of domestic manufacturing firms, and

\[
R^* = p^* z^*(d^*)^{\alpha^*_1} (h^*_t)^{\alpha^*_2} (k^f_t)^{1-\alpha^*_1-\alpha^*_2},
\]

the total revenue from the export sector. Then if we let \(G\) be government expenditure including non-productive purchase, infrastructure investment, and cash transfers, the government budget constraint is

\[
G + \mu^f \tau^r \delta k^f_t = \tau^a(p^a C^a_t + C^m_t) + \mu^f \tau^r (\pi^f_t + \pi^*_t) + \tau^r y^m + \tau^w (\mu^u w^m h^u + \mu^r w^f h^r),
\]

where \(\mu^f \tau^r \delta k^f_t\) is the tax deduction to agricultural machinery investment.

### E. Stationary Equilibrium

We define formally the recursive competitive equilibrium in the steady state.

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

(i) Given \(p, V^j(b^j, \varepsilon^j)\) and \(\{c^x^j(b^j, \varepsilon^j), b^j_t, h^j_t, \varepsilon^j_t\}\) solve the households’ optimization problems (3) and (4);

(ii) Given \(p, \{c^f, k^f, h^r, h^*\}\) solve the large farmer’s optimization problem (5);
(iii) Given \( p, \{ k^m, h^m \} \) solve the firms’ optimization problem (6);

(iv) (Aggregate Consistency) \( \Gamma_j(b^j, \varepsilon^j)s \) are stationary distributions corresponding to the joint transition matrices \( \Pi_j \) constructed from \( b^j(b^j, \varepsilon^j) \) and the transition matrices of \( \varepsilon^j, j = u, r \):

\[
\Pi_j = \Pr[b_{t+1} = b', \varepsilon_{t+1} = \varepsilon'|b_t = b, \varepsilon_t = \varepsilon] = \Pr[b_{t+1} = b'|b_t = b, \varepsilon_t = \varepsilon] \Pr[\varepsilon_{t+1} = \varepsilon'|\varepsilon_t = \varepsilon],
\]

where we suppressed the dependence of \( j \) for simplicity;

(v) Government budget (7) is balanced;

(vi) Prices \( p \) clear all markets:

- **Urban Labor Market:**
  \[
  \mu^u \int \varepsilon^uh^ud\Gamma^u(b^u, \varepsilon^u) = h^m.
  \]

- **Rural Labor Market:**
  \[
  \mu^r \int \varepsilon^rh^rd\Gamma^r(b^r, \varepsilon^r) = \mu^f(h^o + h^s).
  \]

- **Capital Market:**
  \[
  \mu^u \int b^ud\Gamma^u(b^u, \varepsilon^u) + \mu^r \int b^rd\Gamma^r(b^r, \varepsilon^r) = k^m.
  \]

- **Food:**
  \[
  C^a = \mu^r \int z^a \varepsilon^r(d^r)^{\alpha^r}(1 - h^r)^{-\alpha^r}d\Gamma^r(b^r, \varepsilon^r) + \mu^f z^a(d^a)^{\alpha^a}(1 - h^a)^{-\alpha^a}.
  \]

- **Services:**
  \[
  C^s = \mu^u \int z^a \varepsilon^u(d^u)^{\alpha^u}d\Gamma^u(b^u, \varepsilon^u).
  \]

- **Manufacturing Goods:**
  \[
  C^m + \delta(k^m + \mu^f k^f) + G = z^m(k^m)^{\alpha^m}(h^m)^{1-\alpha^m} + \mu^f R^s.
  \]

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.
II. 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 calibrate. In the first group are those parameters whose values we take exogenously from the literature. These parameters are mostly those 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 in order.

**Exogenous Parameters.**—We set the discount rate $\beta$ and depreciation rate $\delta$ 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 in Adamopoulos and Restuccia (2014), leading to 0.49 and 0.32 for $\alpha_1$ and $\alpha_2$ 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 of 0.51. This is consistent with the farming practice of subsistent farming, for which the mechanization level is extremely low. The relative farm size (measured in land) of food to cash crops is set to 0.27 according to Teshome et al. (2016), with the farm size of cash crops normalized to one. Notice that land size will not affect any of our results since they enter the production multiplicatively with sectoral TFP as $z^a(d^r)^{\alpha_1}$ and $z^*(d^*)^{\alpha_2}$. As a result, different values of $d^r$ and $d^*$ would only yield different values for $z^a$ and $z^*$. We include explicitly land in the production to be consistent with the functional forms in Adamopoulos and Restuccia (2014). We assume that the capital and labor shares of the manufacturing production function are the same as commonly used for the United States, 0.63 and 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 productivities, leaving the main results intact. 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 will not be affected as long as this number is small. 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.

**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 match 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 productivities $z^a$, $z^m$ and $z^*$ to match the sectoral shares in output reported by National Bank of Ethiopia. The relative shares for agricultural, manufacturing, and export sectors are respectively 0.42, 0.33, and 0.08. In calculating these statistics, we map the manufacturing sector in our model to include

---

11For readers who are not familiar with the calibration exercise, please refer to Appendix B for more details.

12As specified previously, an alternative choice would be to model the exporting firms as mirroring the zero-profits neoclassical manufacturing firms.
TABLE 1—EXOGENOUSLY CALIBRATED PARAMETERS

<table>
<thead>
<tr>
<th>Variables</th>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount Rate</td>
<td>$\beta$</td>
<td>0.96</td>
</tr>
<tr>
<td>Depreciation Rate</td>
<td>$\delta$</td>
<td>0.06</td>
</tr>
<tr>
<td>Land Share in Food Production</td>
<td>$\alpha^0$</td>
<td>0.49</td>
</tr>
<tr>
<td>Labor Share in Services Production</td>
<td>$1 - \alpha^s$</td>
<td>0.37</td>
</tr>
<tr>
<td>Capital Share in Manufacturing Production</td>
<td>$\alpha^m$</td>
<td>0.37</td>
</tr>
<tr>
<td>Land Share in Cash Crop Production</td>
<td>$\alpha^1_1$</td>
<td>0.49</td>
</tr>
<tr>
<td>Labor share in Cash Crop Production</td>
<td>$\alpha^2_2$</td>
<td>0.32</td>
</tr>
<tr>
<td>Relative Land Size of Food Farm</td>
<td>$d^r$</td>
<td>0.27</td>
</tr>
<tr>
<td>Persistence of Urban Income Shocks</td>
<td>$\rho^u$</td>
<td>0.92</td>
</tr>
<tr>
<td>Persistence of Rural Income Shocks</td>
<td>$\rho^r$</td>
<td>0.92</td>
</tr>
<tr>
<td>Urban Population Share</td>
<td>$\mu^u$</td>
<td>0.28</td>
</tr>
<tr>
<td>Rural Population Share</td>
<td>$\mu^r$</td>
<td>0.69</td>
</tr>
<tr>
<td>Large Farmer Share</td>
<td>$\mu^f$</td>
<td>0.03</td>
</tr>
</tbody>
</table>

modern services industry such as telecommunication, banking, etc., as they resemble manufacturing firms operating on the formal market. Similarly, the subsistence level $\bar{a}$ and relative preference $\gamma$ and $\psi$ are calibrated to match the food share by the bottom quintile people and relative consumption shares 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.

Parameterizing the Income Processes.—The income processes warrant more explanation. In theory, there are four AR(1) processes, each region has two, one for formal and one for informal market:

$$
\begin{align*}
\varepsilon^{u}_{t+1} &= \rho^u \varepsilon^{u}_{t} + \eta^{u}_{t+1}, \\
\varepsilon^{r}_{t+1} &= \rho^r \varepsilon^{r}_{t} + \eta^{r}_{t+1}, j = u, r.
\end{align*}
$$

The data requirement for estimating the above processes are very demanding. Ideally, we would need to observe the income from both formal and informal sources for a same household, and in addition a panel structure. 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 later. In particular, we first assume that the formal and informal shocks are correlated in the sense that one is constructed from the other according to the details specified as follows. This reduces the number of processes we are dealing with to two, one for each region, which we call the root processes. We then assume that both processes have a $\rho$ equals 0.92, which is slightly lower than the commonly used value for the United States 0.95 [Storesletten, Telmer and Yaron (2004a,b)].\textsuperscript{13} Let the variances of the innovations in urban and rural areas be $\sigma^2_u$ and $\sigma^2_r$. For any values of $\sigma^2_u$ and $\sigma^2_r$, we can use the

\textsuperscript{13}The literature has shown that the approximation on the variance of the stationary distribution deteriorates as $\rho$ approaches 1 for both the Tauchen (1986) and Tauchen and Hussey (1991) methods, see Flodén (2008) for details. We choose a slightly smaller $\rho$ here for computational concern. Our sensitivity analysis over $\rho$ shows that it does not affect our results.
<table>
<thead>
<tr>
<th>Data Targets</th>
<th>Parameters</th>
<th>Value</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food Share in Consumption by Bottom Quintile</td>
<td>$\pi$</td>
<td>0.02</td>
<td>0.52</td>
<td>0.50</td>
</tr>
<tr>
<td>Manufacturing Share in Consumption</td>
<td>$\gamma$</td>
<td>0.82</td>
<td>0.33</td>
<td>0.35</td>
</tr>
<tr>
<td>Services Share in Consumption</td>
<td>$\psi$</td>
<td>0.50</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>Rural Consumption Gini</td>
<td>$\sigma_r^2$</td>
<td>0.23</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>Urban Consumption Gini</td>
<td>$\sigma_u^2$</td>
<td>0.63</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>Tax to GDP Ratio</td>
<td>$\tau^a$</td>
<td>0.06</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>CIT in Total Tax Revenues</td>
<td>$\tau^r$</td>
<td>0.12</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>PIT in Total Tax Revenues</td>
<td>$\tau^w$</td>
<td>0.06</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>Food Share in Output</td>
<td>$z^a$</td>
<td>0.74</td>
<td>0.42</td>
<td>0.34</td>
</tr>
<tr>
<td>Manufacturing Share in Output</td>
<td>$z^m$</td>
<td>10.19</td>
<td>0.33</td>
<td>0.38</td>
</tr>
<tr>
<td>Export Share in Output</td>
<td>$z^*$</td>
<td>0.68</td>
<td>0.08</td>
<td>0.10</td>
</tr>
</tbody>
</table>

standard Tauchen (1986)'s method to approximate these processes using Markov Chains. We take the process of the formal market as the root in the urban area, and that of the informal as the root in the rural area. We then “inflate” the shocks of the root processes that are less than the mean by taking a geometric average of the corresponding shocks with the maximum, using a weight of 0.95 for the maximum. This lets urban households that are less productive have a comparative advantage in the informal market, while the less productive rural households gain an advantage in the formal market, and hence creates the sorting of workers documented by Eastwood, Lipton and Newell (2010) and La Porta and Shleifer (2014) mentioned earlier. We then calibrate $\sigma_u^2$ and $\sigma_r^2$ to match the observed consumption Ginis in the rural (0.26) and urban areas (0.40), both retrieved from the World Bank. Table 2 lists the values used for the endogenously calibrated parameters and the model fit. Overall, the model does a good job to fit the data, except that manufacturing goods and cash crops are slightly overshot, while food is undershot.

Discussion.—Though it may seem that we have taken some liberty in parameterizing the income processes, we argue that for the purpose of our paper, these liberties would not affect our results for two reasons. First, for a given spread of income shocks, the persistence level primarily affects the risk of income, and hence the precautionary motives. Since for most low income countries, households save little of their income, the quantitative impacts of risks are small. In fact, in our benchmark calibration, the national savings rate is just above 10% with a high persistence of 0.91. The savings rate is even smaller in the rural area, where 70% of the population live. Since savings are by assumption bounded from below by zero, we believe that the choice of $\rho$ will not affect our quantitative results. Second, the geometric weights affect mostly how households divide their hours to formal and informal market. Because these incentives are also affected by the sectoral TFP, and we are calibrating the sectoral productivity to the production shares in output, a different (for instance, smaller) weight in the urban area would just lead to a smaller calibrated $z^m$, and would not affect the results. Nonetheless, in the sensitivity analysis, we verify the robustness of our results by solving the model under different $\rho$s and geometric weights.
III. Quantitative Results

We use the calibrated model to compare the macroeconomic and distributional impacts of different revenue mobilization schemes and fiscal expenditure policies. We first compare the effects from raising tax revenues equaling 2% GDP in the benchmark steady state through VAT, PIT, and CIT. In all three experiments, the tax revenues are assumed to be used as government expenditure not valued directly by households. The idea of these experiments is to investigate that if the country authority has to raise a certain amount of tax revenue, which of the three instruments is least costly. We then investigate the effects of using the additional tax revenue for universal basic income (UBI) and public infrastructure investment that boosts the overall TFP of the economy. At last, we examine the robustness of our results to certain modeling assumptions. Specifically, we run sensitivity analyses over the persistence of income shocks $\rho$, the curvature parameters that determine the sorting of labor, and raising tax revenues equaling 5% GDP in the benchmark steady state. In the interests of space, the sensitivity results are not reported, but they are available upon requests.

A toolkit with a Graphical User Interface along with the parameterization files of all the simulations is provided to the users. Appendix C provides more details on how to obtain the toolkit and replicate the results in the paper, and to use the model for future policy analyses.

A. A Comparison between Tax Instruments

In the first set of experiments, we raise tax revenue of 2% current GDP by each of the three tax instruments, and let the government use all the revenue to fund non-productive government expenditure on manufacturing goods.

Macroeconomic Impacts.—In Figure 1, we show the resulting percentage change from the benchmark steady state of six moments caused by the three tax instruments. From left to right, the six moments reported are aggregate output, aggregate consumption, aggregate investment, urban, rural, and total consumption Gini coefficients. We focus on consumption Gini here because our model is calibrated to consumption inequality. However, income Ginis reserve the same relative order of the impacts from different taxes, and thus deliver the same message.\(^{14}\) In addition, because in this family of models most of the time taxes are distortionary, it is not surprising that all taxes lead to output decline.\(^{15}\) Therefore, similarly, what matters more here is the ranking of the relative impacts of different taxes, rather than the actual increase or decrease.

We find that the efficiency costs measured by output and consumption losses are largest for PIT, with CIT and VAT following in order. The result that VAT is least distortive follows immediately from the fact that if $\bar{\alpha} = 0$ and we can tax services goods at the same rate, then VAT will not lead to

---

\(^{14}\)For instance, the impacts of VAT, PIT, and CIT on total consumption Ginis are 2.26%, −0.48%, and −0.78%. Their impacts on total income Ginis are 1.91%, 1.41%, and −0.38%. Though the effect of PIT on total income Gini is positive and on total consumption Gini is negative, the ranking of the effects from the three taxes is VAT > PIT > CIT for both cases.

\(^{15}\)There are some exceptions of course. 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 distortionary. The Stone-Geary preference in our model violates the second condition.
output loss. Since in the benchmark calibration, $\alpha = 0.02$ which is very close to zero, and services goods expenditure only takes 22% of household’s budget on average, the resulting distortions are small. PIT is more distorting than CIT because it directly discourages households from working in the sector that they have comparative advantage in, while the impacts of CIT come through second order effects of prices.

The increase in the taxes have differential impacts on sectoral output. From the top panel of Table 3, three observations stand out. First, manufacturing sector is always the least affected across all experiments. In fact, for the VAT scenario, it even increases by 1.32%. This is due to the fact that all tax revenue is spent to purchase manufacturing goods by the government, thereby creating extra demands. If instead we distribute the additional tax revenues back to the households as in the experiments in the next section, the drop in manufacturing output is larger than agriculture and services, due to decreasing relative prices of food and services caused by reduction in average household demand (Table 4). Second, the decrease in agricultural and services output are of similar magnitudes for the cases of PIT and CIT, while for the VAT case, services are less affected. This is straight forward to understand, as VAT directly increase the relative price of food and manufacturing goods to services. In addition, the similar magnitudes of impacts across consumption and production suggest that flattened consumption drives most of the changes in macro aggregates. Third, the fact that the decreases in consumption of the three goods are roughly of the same magni-

---

16 Consumption will drop however, since what is kept constant is household’s consumption expenditure, for which consumption tax takes a bite.
Table 3—Sectoral and Regional Impacts: Baseline

<table>
<thead>
<tr>
<th></th>
<th>VAT</th>
<th>PIT</th>
<th>CIT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Moments</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sectoral Output</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>−5.41%</td>
<td>−9.50%</td>
<td>−7.55%</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>1.32%</td>
<td>−3.42%</td>
<td>−2.60%</td>
</tr>
<tr>
<td>Services</td>
<td>−2.04%</td>
<td>−9.50%</td>
<td>−7.55%</td>
</tr>
<tr>
<td><strong>Rural Consumption</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>−8.06%</td>
<td>−9.32%</td>
<td>−6.88%</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>−8.07%</td>
<td>−9.31%</td>
<td>−6.88%</td>
</tr>
<tr>
<td>Services</td>
<td>−4.79%</td>
<td>−9.31%</td>
<td>−6.88%</td>
</tr>
<tr>
<td><strong>Urban Consumption</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>−3.48%</td>
<td>−10.72%</td>
<td>−7.36%</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>−3.48%</td>
<td>−10.72%</td>
<td>−7.36%</td>
</tr>
<tr>
<td>Services</td>
<td>−0.03%</td>
<td>−10.72%</td>
<td>−7.36%</td>
</tr>
</tbody>
</table>

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

The table implies that in our calibrated economy, the consumption expenditure of most households are higher than subsistence requirement.

**Distributional Impacts.**—The distributional property of the economy comes from two sources: within region inequality and between region inequality. The three taxes have different distributional implications along these two dimensions.

The results are presented in the lower two panels of Table 3. We first find that between regions, rural households are more affected in percentage than urban households for a VAT increase, but the impacts are reversed for PIT and CIT. Furthermore, for VAT, it is not just because rural households on average consume less than urban households, hence a drop of consumption of the same absolute level would mechanically lead to higher percentage decline. The actual decline of consumption in level is higher in the rural area. This is due to two reasons. One comes from the general equilibrium effects of lowered food price due to depressed private consumption, which disproportionately reduce the income of rural households. The other is because for Stone-Geary preference, the risk aversion of households increase with increasing subsistence requirement and decreasing disposable income [Chetty and Szeidl (2007)]. The reason that for PIT and CIT, urban population is affected more is because on average the tax burden from these two taxes is mostly imposed on the urban population. In particular, both PIT and CIT only affect the income from the formal markets directly. A Ramsey tax thus implies that the higher taxable income, the more tax revenue collected from that source. Since in the urban area more productive households supply more hours to the formal market, while the opposite is true for the rural area, urban households are disproportionately taxed heavier. These observations when combined, explain the reason that in Figure 1, total consumption Gini increases for the case of VAT, while decreases for PIT and CIT. However we notice that the magnitudes for PIT and CIT are quantitatively small.

Recall that for preference (1), changes in consumption expenditure on each goods are symmetric when $\pi \rightarrow 0$. In particular, this is due to the logarithmic assumption which cancels the substitution and income effects. A more general CES functional form would lead to changes that are in proportion to each other.
FIGURE 2. CONSUMPTION CHANGES BY DECILES: BASELINE

Note: The results show the average consumption by deciles with respect to the benchmark case.
The within region impacts are more complicated. The changes in average consumption across different deciles for rural and urban households are shown respectively for the three taxes on the left and right columns of Figure 2. We adjust the y-axis of the figures for demonstration purposes, so these figures are plot on different scales. Among the two areas, the patterns from the urban area are easier to explain. As is suggested from the changes in the Ginis in Figure 1, VAT is regressive, PIT is progressive, while CIT has virtually uniform impacts. These aggregate figures are consistent with the distributional impacts from the right column of Figure 2. Theoretically, there are three mechanisms undergoing here. First is the changing marginal propensity to consume due to the concave consumption function. Second is the tax incidence. And third, there is the subsistence requirement that prevents households from further reducing their consumption. For VAT, poor households are affected more because they have less savings to buffer consumption fluctuation. For PIT, on the other hand, rich households cut expenditure more because they are taxed disproportionately heavier under the scenario. The impacts of CIT come in principle from the general equilibrium changes in prices, and hence are evenly distributed across households. The urban households are slightly more affected, because firms would transfer a fraction of the tax burden through interest rates and formal wages, both would have larger impacts to rich households.

The rural patterns are more convoluted. Qualitatively, the impacts of VAT and CIT across the consumption distribution are similar. Average consumption decreases more as households become richer until the third decile, while the opposite is true for households above the fourth decile, making the pattern overall U-shaped. The non-monotonicity occurs around the third decile is caused by a switch of occupation of households from relying more on wage income as hired labor to more on sales of food. For the ease of reference, we label them as tenants and family farmers. The distributional impacts of VAT and CIT on family farmers are similar to those on urban households, and hence the same rationale applies. For tenants, because they are those with the lowest income, the subsistence requirement prevents them from further cutting consumption, which leads to lower percentage drop in consumption for poorer tenants. PIT on the other hand, reverses the whole pattern to an inverse-U shape. Like the cases of VAT and CIT, family farmers’ behavior still resembles those of urban households. Here poorer tenants cut their consumption more because of income effect. Recall that PIT are raised mostly on tenants, whose major source of income is formal rural labor market. Under the labor sorting assumption, tenants are disproportionately taxed by PIT, hence the decrease in their incomes forces them to reduce consumption, even if they are already close to the subsistence level. The mapping of changes in consumption by deciles to changes in Gini coefficients are less transparent for the rural area. Though Figure 1 suggests that all three taxes reduce inequality in the rural area, results in Figure 2 clearly show that a lot of variations are masked. In fact, in close scrutiny, we find that PIT could be astoundingly regressive within certain deciles and damage households in extreme poverty the most.

Overall, we find that among the three taxes, VAT incurs the lowest efficiency costs in terms of aggregate output and consumption, but it could be very regressive by widening an already large rural-urban discrepancy in low-income countries [Young (2013)]. CIT, on the other hand, though causes larger efficiency costs, but has considerable better inequality implications. PIT, however,

\[\text{We use these terminology purely for presentation purpose, readers should not connect them with more nuanced micro institutional arrangements.}\]


### Table 4—Sectoral and Regional Impacts: Different Policies

<table>
<thead>
<tr>
<th>Moments</th>
<th>Rural UBI</th>
<th>High UBI</th>
<th>Infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sectoral Output</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>1.88%</td>
<td>1.79%</td>
<td>3.54%</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>−1.50%</td>
<td>−1.80%</td>
<td>−3.64%</td>
</tr>
<tr>
<td>Services</td>
<td>1.88%</td>
<td>1.80%</td>
<td>3.54%</td>
</tr>
<tr>
<td><strong>Rural Consumption</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>4.52%</td>
<td>3.65%</td>
<td>7.26%</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>4.53%</td>
<td>3.66%</td>
<td>7.27%</td>
</tr>
<tr>
<td>Services</td>
<td>4.54%</td>
<td>3.66%</td>
<td>7.27%</td>
</tr>
<tr>
<td><strong>Urban Consumption</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>−0.04%</td>
<td>0.52%</td>
<td>0.99%</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>−0.04%</td>
<td>0.52%</td>
<td>0.98%</td>
</tr>
<tr>
<td>Services</td>
<td>−0.04%</td>
<td>0.52%</td>
<td>0.98%</td>
</tr>
</tbody>
</table>

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

deteriorates both the economic efficiency and equity, thus is the most detrimental instrument.

### B. Cash Transfers and Infrastructure Investment

Inspired by the discussions above, in this section we evaluate whether more productive use of the tax revenue for instance on cash transfers or public investment could mitigate the negative impacts of fiscal mobilization. Throughout this section, we assume that the new benchmark is the one where the government raises tax revenue equals to 2% GDP using VAT. We then consider four experiments: half of the additional tax revenues are used for cash transfer to rural households only (case 1), and for universal basic income (UBI, case 2), all additional tax revenue is used for UBI (case 3), and half of the tax revenue is used to finance infrastructure investment that boosts economy-wide TFP (case 4). We focus on VAT for two reasons. The first is that VAT has the smallest costs on efficiency as is shown in the previous section. Hence a mitigation of its negative inequality impacts could make it an attractive fiscal consolidation instrument when packaged together with cash transfer programs. Because the negative distributional impacts of VAT are mostly on the between region inequality, we investigate specially the case of cash transfer to rural households only. The second is because the additional impacts of different expenditure programs are similar regardless of the revenue sources, hence it suffices to analyze VAT as a representative case. Notice that the total impacts of the fiscal packages considered in this section comparing to the original benchmark economy can be inferred from adding the results in this section to those in the previous section.

**Cash Transfer Programs.**—The impacts of different cash transfer programs are summarized in Figure 3. 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 cash transfers. In particular, we find that case 1 leads to higher increase in output and consumption, less drop in private investment, and overall larger improvement in inequality. These findings confirm several economic intuitions in the last section. First, since
most of the worsening in inequality under the VAT case is caused by between regions inequality, compensation for the poor rural households naturally reduce the overall inequality. Second, from Table 4, we find that rural households increase their consumption more than urban households in response to cash transfers due to their higher marginal propensity of consumption [Carroll (2009)]. This also explains the decline of within region Gini coefficient. We show the changes of consumption by deciles for case 1 in Figure 4. Here we refrain from presenting those from cases 2 and 3, as qualitatively they are almost the same as case 1. The monotonic decrease of consumption along the consumption distribution is generated by general equilibrium price effects. Specifically, with cash transfers, extra demand from rural households drive up food and services prices and suppresses manufacturing wage, which increases the income of poor urban households relying on informal services and decreases that of the rich who rely on formal wage. Third, the decrease in investment (Figure 3) and manufacturing output (Table 4) suggest that the increase in investment and manufacturing output in previous cases come indeed from the concentrated government expenditure on manufacturing goods. When the expenditure is dispersed as is shown in the bottom two panels of Table 4, manufacturing output decline as food and services.

The comparison of cases 2 and 3 illustrates the effects of a uniform expansion of cash transfer programs, which naturally leads to higher output and lower inequality as shown in Figure 3.

Overall, a cash transfer program targeting rural households with a size of 1% GDP is able to completely overturn the negative inequality impacts of VAT, and brings back about 40% of the total consumption drop due to additional tax liability. A cash transfer program of the same budget
but instead targets the whole population has similar qualitative impacts on the economy, but is less efficient since rural households are more affected by taxes. Though further expansion of cash transfer programs could continue pull up consumption and suppress inequality, using too much tax revenues for redistribution to some extent defeats the purpose of fiscal consolidation, which is to improve the fiscal buffer of the government.

**Infrastructure Investment.**—We model the effects of public investment as boosts of TFP in all sectors in a reduced-form way. Specifically, if we let $K_g$ be the level of public investment, then sectoral productivities are increased to

$$z_j = z_j \exp(K_g)^{\alpha_g}, \ j = a, m, s, \ast,$$

where $\alpha_g$ controls the overall efficiency of public investment.\(^{19}\) Because higher $\alpha_g$ mechanically boosts the productivity more, and most low-income countries are unlikely to have high efficiency in public investment, in our exercise we choose a conservative value of 0.05.\(^{20}\) In our experiment, public investment of 1% GDP translates to a 1.74% increase in overall TFP. The results are shown together with those from the two equal size cash transfers programs (repeated from Figure 3) in Figure 5. As expected, public investment generates higher output and consumption, it also spurs more private investment due to elevated returns. The negative impacts of VAT on output is completely offset from a 1.49% loss to 1.11% increase. Similarly, 54% of the loss in consumption is restored. However, because sectoral productivities are boosted proportionately, the states of inequality in the economy barely changes regardless of measurement by Gini coefficients or consumption distribution. As a result, though public investment can make up for most of the efficiency costs, it is silent on redistributional concerns.

\(^{19}\)Notice that here $K_g$ is the level of investment as opposed to stock. We use $\alpha_g$ to capture the fact that a one billion dollar public investment usually has different effects for countries with different initial infrastructure stock. The exponential guarantees that zero investment does not lead to productivity improvement.

\(^{20}\)Estimates by Berg et al. (2013) indicate that the elasticity is usually between 0.1 and 0.2.
Note: We truncate the vertical axis at 5% to enhance the visual presentation of other variables. Investment, which includes public investment, in fact increases by 17.8%. All results are percentage change with respect to the VAT case.

IV. Conclusion

In this paper, we quantitatively investigate the macroeconomic and distributional impacts of fiscal consolidations in low-income countries through VAT, PIT, and CIT. We find that VAT has the least efficiency costs but is highly regressive, while PIT and CIT lead to higher output and consumption drop, but have better distributional implications. Further, we find that cash transfers targeting rural households are able to mitigate the negative distributional impacts of VAT, while public investment shows almost no distributional impacts.

Our results suggest that low-income countries indeed face very different equity-efficiency trade-offs comparing to advanced economies due to their unique economic structure. It therefore is important to investigate quantitatively the impacts of other tax instruments that have been extensively studied under the environment of advanced economies. For instance, we believe that the optimal progressivity of income tax is also a critically important feature in low-income countries. It is also interesting to study whether and how will classical optimal taxation results under complete markets change when migrated to an economy resembling low-income countries. Another limitation of our study is the omission of transitional dynamics in our model, which is important to the evaluation of short-run welfare effects. We leave these extensions to future work.
Appendix

A. Computational Algorithm

The model is solved by discretization. To proceed, we denote the asset space by $B$, and the idiosyncratic shocks spaces as $E_r, E_u$ 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 $B = \{b_1, \ldots, b_n\} \subseteq B$.\footnote{Geometric grids puts more grids when $b$ is small, since according to the consumption saving literature [Aiyagari (1994) and Carroll (1997)], this region where the policy functions are likely to be nonlinear.}
   (ii) Approximate the shock distributions of productivity in rural informal and urban formal sectors by both types of households using Tauchen (1986)’s method. Denote the $m$-states Markov chain by $\varepsilon_x = \{\varepsilon_{1,x}, \ldots, \varepsilon_{m,x}\} \subseteq E_x, x = \{r, u\}$.
   (iii) For urban households, set the formal sector productivity to be $\tilde{\varepsilon}_{j,u} = \min\{\varepsilon_{m,u}^{\nu}, 1\}$, $\nu \in (0, 1)$. This “inflates” the informal productivity for less efficient urban workers (less than one). As a result, less efficient urban workers choose to work in the services sector (informal) more. Similarly, recall that rural formal productivity is normalized to one. This leads to less efficient rural households spend more time working for large farms (formal).

2. Make an initial guess of the price vector $P_0 = \{p^a, p^s, w^m, w^f, r\}$.

3. Given $P_0$, solve the recursive optimization problems of the urban and rural households (3) and (4). 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

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

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 Piecewise Cubic Hermite Interpolation (PCHI).

   (i) For each node $(b_i, \varepsilon_j)$, 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 both the Hermite polynomial and its derivative. 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 $(b_i, \varepsilon_j)$.

(v) Loop over all nodes in $B$ and $\varepsilon_u$ 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 step (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. Recall that in steady state, it has analytical solution for given prices.

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 $B$. In particular, we construct linear grids $\overline{B}_1 = \{b_1, \ldots, b_{n_1}\}$, 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 $(b_i, \varepsilon_j)$, 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 in $\overline{B}_1$. Under such scenarios, we first locate the interval in $\overline{B}_1$ that contains $b'(b_i, \varepsilon_j)$, we then distribute the share of people to the endpoints of the interval based on the distance between $b'(b_i, \varepsilon_j)$ and the endpoints. However, $\varepsilon'_j$ will always be on $\varepsilon_x$, owing to the discrete nature of the Markov Chain.

(iv) Repeat (iii), until the two successive density functions converge, where by stationarity the density function is 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 [see Section 10.7.2 of Press et al. (2007)] to update the guess of $P$. 
The above algorithm solves the steady state equilibrium of the model under *given* parameterizations. To calibrate the model, we would choose the parameterizations of the model such that model implied moments match those in the data under certain measure of distance.\textsuperscript{22}

In practice, the model is parallelized using MPI, where the execution of the code is paralleled over the shock states $\varepsilon_j$ when solving the Bellman Equations (step 3).

B. Mathematical Foundation of Calibration

*The Solution Concept.*—In a competitive economy, all agents (including households, firms, farmers, government) take prices as given. Therefore, for any vector of prices, the solutions to the optimization problems characterize the allocation of resources across different agents. These individual level allocations can then be aggregated to calculate values of the aggregate variables. Therefore, in a narrow sense, the solution of the model is simply a vector of prices that clears all the markets. We call this particular vector of prices the *equilibrium price*. All allocations can be solved accordingly given the equilibrium price. For reasons that should be obvious, if the value of any parameter of the model is changed, a new equilibrium price has to be solved. Furthermore, if you know the equilibrium price vector, the allocations that are solved under this price vector, will by definition clear all the markets. In what follows, we will only consider solution of the model under equilibrium prices.

*Calibration.*—We use macroeconomic models to study various aspects of the economy as a whole, for instance real GDP per capita, Gini coefficient, investment to output ratio, etc. The underlying concept is that we would like to use the model as a “measuring device,” to measure the effects of changes in certain model parameters that parsimoniously capture some policy changes in reality. As a result, we would like to work with a solution that replicates salient features of the actual economy. Often this is done by comparing the model implied values of certain economic moments along several key dimensions with those found in the data. This parameterization of the model that replicates stylized facts of the economy under current policy is usually called the *benchmark*. For most well-defined macroeconomic models, one particular parameterization of the model corresponds to a particular solution, characterized by a vector of values for model implied statistics. For instance, if $\delta = 0.1$, the capital to output ratio implied by the model might be 3.3, while if $\delta = 0.05$, it might be 2.0. Therefore, if the actual capital to output ratio in the data is 3.0, we would want to set $\delta$ to 0.1 as opposed to 0.05. Put in plain English, from a mathematical point of view, calibration is to choose the parameterization of the model, such that a certain distance between the model implied moments and their data counterparts is minimized. In essence, it is a multi-dimensional unconstrained optimization problem.\textsuperscript{23} We then change some parameters of the model to reflect features of the policy that we would like to implement, for instance a 3 percentage points increase in VAT. The solution of the model under this new parameterization yields a new set of prices and allocations. Because the only difference between the new model and the benchmark model is the values of the parameters that you change, all changes in the characterization of

\textsuperscript{22}See Appendix B for details.

\textsuperscript{23}For a formal elaboration of the calibration process, please refer to the handbook chapter by Dawkins, Srinivasan and Whalley (2001).
the equilibrium must be caused by the changes of the values of these parameters. Readers who are familiar with modern applied micro-econometrics could think of the exercise as identifying the causal effects of some mechanisms with the help of a model.

**Calibration as a Minimization Problem.**—The minimization problem characterizing the calibration process is formally illustrated as follows. Throughout, we assume that all vectors are column vectors. A macroeconomic model can be viewed as a non-linear operator that maps a group of parameters to values of certain moments. Let the spaces of parameters and moments to be \( \mathcal{P} \) and \( \mathcal{S} \) respectively. The model is thus an operator \( \mathcal{F} : \mathcal{P} \to \mathcal{S} \). Under this notation, for a vector \( p \in \mathcal{P} \), from the solution of the model (which we define here as the stationary equilibrium), we can calculate the values of a list of moments \( \mathcal{F}(p) = s \). In the context of the current paper, \( p \) is an eleven-dimensional vector \( (a, \psi, \gamma, \cdots, z^*)' \), and \( s \) is another eleven-dimensional vector where \( s_1 \) is the model implied food share of the bottom decile households in total consumption, and \( s_{11} \) is the model implied share of exports in total output. With the default parameterization provided by the toolkit, we have \( p = (0.0208, 0.4945, 0.8168, \cdots, 0.6845)' \), and \( s = (0.5010, 0.2227, 0.3489, \cdots, 0.1678)' \). In addition, for every moment represented by \( s \), we can find their data counterparts, which we denote by a vector \( d \). In the context of the current model, \( d_1 \) and \( d_{11} \) are the food share in consumption by bottom decile households and the share of exports in total output found in the data. In our paper, \( d = (0.52, 0.21, 0.33, \cdots, 0.08)' \). With the notations defined so far, the mathematical problem of calibration can be characterized as the following quadratic optimization problem.

**Definition 2.** (Calibration) Let \( \mathcal{P} \) be the space of all feasible parameter values, and \( \mathcal{S} \) be that of all possible values of some moments of interest. A model defines an operator \( \mathcal{F} : \mathcal{P} \to \mathcal{S} \). If we let the data counterparts of moments in \( \mathcal{S} \) be \( d \), then calibration is a quadratic minimization problem

\[
(B.1) \quad p^* = \arg\min_{p \in \mathcal{P}} \{ [\mathcal{F}(p) - d]' W [\mathcal{F}(p) - d] \},
\]

where \( W \) is a given weighting matrix.

If we assume \( \mathcal{P} = \mathbb{R}^n \), then (B.1) is simply an unconstrained optimization problem. If we let

\[
\mathcal{G}(p) \triangleq [\mathcal{F}(p) - d]' W [\mathcal{F}(p) - d],
\]

then standard results from mathematical analysis establishes that \( p^* \) is a local minimizer to (B.1), if and only if

\[
\nabla \mathcal{G}(p^*) = 0,
\]

and the Hessian Matrix of \( \mathcal{G} \) at \( p^* \), \( D^2 \mathcal{G}(p^*) \) is positive definite.

The minimization problem (B.1) is mathematically challenging. Besides the complication that the first and second order conditions only guarantee local solution, even searching for such a local solution proves to be a daunting task. More specifically, although \( \mathcal{G}(p) \) is quadratic in \( \mathcal{F}(p) \), in most cases it is not quadratic in \( p \). Take the scalar case as an example, when \( p \) is a scalar, the minimization problem (B.1) becomes

\[
p^* = \arg\min_{p \in \mathcal{P}} \mathcal{G}(p).
\]
The first and second order conditions
\[
\frac{dG(p^*)}{dp} = 0, \quad \frac{d^2G(p^*)}{dp^2} > 0,
\]
usually do not hold.

C. Use the Toolkit to Replicate the Results

The toolkit can be obtained from the Knowledge Exchange (KE) site through Fund’s Intranet at: http://intranetapps.imf.org/fundwide/ke/Topics/Income-Inequality/Pages/toolkit.aspx

Alternatively, readers outside the Fund could email the authors to request for a copy of the toolkit. Figure A1 shows a screenshot of the KE site hosting the toolkit. The toolkit is distributed as a zip package. Complete documentation of the toolkit is provided under the .//Documentation/ folder of the decompressed package. The users are highly recommended to read through the Release Notes and the Tutorial before start using the toolkit. To run the toolkit, a MATLAB of version higher than 2014b is required. The toolkit at this moment only runs on Microsoft Windows system of version later than Vista. To accelerate the execution by paralleling, Intel MPI Library Runtimes are also required.

Figure A2 shows the main Graphical User Interface of the toolkit. The readers can use the Load feature to load all parameterizations for the experiments in this paper and replicate the results. For more details, please refer to the documentation of the toolkit.
The model assumes that the government runs a balanced-budget, meaning that in each period, government expenditure equals tax revenue. However, in reality, most countries accumulate a certain amount of government debt. It therefore creates a discrepancy between the model and the real economy. In this section, we explain qualitatively how the results change if government debt has been explicitly modeled. It would help the readers interpret the simulation results from a model where the government runs balanced-budget calibrated for countries that actually issue public debt. In summary, modeling government debt explicitly would lead to similar response on the output of the economy, but with a different implication on interest rate. The intuition is exactly that of moving from an Aiyagari (1994) model to an Aiyagari and McGrattan (1998) model. The left panel of Figure A3 illustrates the asset market clearing condition of the Aiyagari (1994) model. In the figure, the y-axis is the interest rate and the x-axis is the quantity of assets. The red curve $A(r)$ plots the aggregate saving of households (supply of assets) as an increasing function of the interest rate $r$. Notice that due to market incompleteness, assets supply approximates infinity asymptotically as interest rate approaches the time preference $\lambda$. Similarly, the blue curve $K(r)$ plots the capital demand of firms (demand of assets) as a decreasing function of the interest rate. In equilibrium, asset demand equals asset supply, which pins down the equilibrium interest rate $r^*$. 

**Figure A2. Load Parameterizations Using the GUI**

**D. The Role of Public Debt**
Now we introduce government debt to the model. If we let $G_t, B_t$ and $T_t$ represent respectively government expenditure, debt and tax revenue in period $t$, then the government budget constraint becomes

$$G_t + (1 + r_t)B_t = B_{t+1} + T_t,$$

where if we set $B_{t+1} = B_t = 0$, we return to the usual case of balanced-budget.

We assume that government debt is perfect substitute of private capital $k$, meaning that it is subject to the same interest rate. As a result, households only care about the total amount of saving, but not the composition between capital and government debt. As a result, under the new setting, the aggregate saving function $A(r)$ would be the same as in Figure A3. The asset demand function changes by shifting the original $K(r)$ rightward by $B$. This comes from the fact that $K(r)$ is derived from the first order condition of the firms, which remains the same in the presence of government debt, and government simply competes with private firms for the total savings available in the economy. This is presented in the right panel of Figure A3. Hence in equilibrium, the interest rate increases from $r^*$ to $r_1$, and private capital $K$ decreases accordingly. This is simply the classical crowding out effect of public debt embedded in a heterogenous agents model.

**Example: Debt Repayment.**—As an example, we consider the case where the additional tax revenue is used to repay outstanding debt. This essentially moves the asset demand curve from the dashed line leftward to the solid line in the right panel of Figure A3. In the new equilibrium, the *debt mechanism alone* would lead to a decrease in interest rate and an increase in total output. On the other hand, if the additional tax revenue is used to finance government expenditure as in the model in the main text, then this is equivalent to a *positive demand shock* to firms. As a result, for any given interest rate $r$, facing a positive demand shock, firms would demand more capital. Graphically this shows as a rightward shift of the capital demand curve from $K_0(r)$ to $K_1(r)$. We demonstrate this in Figure A4. In the new equilibrium, both interest rate and output increase (from $r^*$ to $r_1$).

So comparing the cases with and without government debt, we see that qualitative the implications on output are the same. What differentiates the two cases is the model predictions on interest
The fundamental reason is that government debt changes the capital supply curve that firms face, while government expenditure changes the capital demand curve. Since across all experiments, we have very low equilibrium interest rates, at least for the benchmark calibration, whether government debt is modeled explicitly or not would not change the quantitative message of the model.
REFERENCES

Allen, Treb, and David Atkin. 2017. “Volatility and the Gains from Trade.” Manuscript. 6


