

Modeling Distributional Effects of Macroeconomic Shocks: Increasing female participation and human capital in Turkey

Israel Osorio Rodarte¹

Draft not for citation

Washington, D.C. April 15, 2016

Abstract

The Global Income Distribution Dynamics is a top-down macro-micro simulation modeling framework used to generate counterfactual income distributions based on changes in economywide macroeconomic conditions. This paper documents latest innovations in the GIDD model. While it remains focused on the microeconomic underpinnings, this paper is intended to be used by macroeconomic modelers interested in adding an extra distributional modeling layer to simulations. Better coverage of harmonized household surveys permitted a more consistent treatment of macroeconomic inputs and allow for increased detail in the microeconomic model. Macro inputs from a recursive dynamic general equilibrium model and GIDD simulations suggest progressive distributional effects from increased female labor participation and human capital accumulation in Turkey. Possible future extensions to the modeling framework are discussed.

¹ Development Prospects Group, The World Bank. iosoriorodarte@worldbank.org

1 Introduction

The GIDD is a top-down macro-micro simulation framework that distributes macroeconomic shocks exploiting the observed heterogeneity available in household surveys. GIDD was developed by the World Bank and was inspired by previous efforts involving simulation exercises (François Bourguignon, Ferreira, and Leite 2008; François Bourguignon, Bussolo, and Pereira da Silva 2008; Davies 2009). Earlier versions of the GIDD model can be found in François Bourguignon & Bussolo, (2013); and Bussolo, De Hoyos, & Medvedev, (2010).

The GIDD has often been linked with the LINKAGE global computable general equilibrium (CGE) model (van der Mensbrugghe 2011); but it can be linked with other dynamic CGEs or with simpler macroeconomic projections. The type of policy simulations is varied. Recent modeling applications include the effects of demographic change (Ahmed et al. 2014); climate, violence, and global economic stagnation (Devarajan et al. 2015), deeper regional trade integration (Balistreti et al. 2016), or the effects of China’s economic slowdown and rebalancing.

This paper provides a comprehensive documentation of the GIDD mechanics. While this paper remains focused on the microeconomic underpinnings, it is intended to be used by macroeconomic modelers interested in adding an extra distributional modeling layer to simulations. GIDD’s top-down modeling framework allows it to be linked with macroeconomic scenarios and forecasts obtained from a variety of third-party models. Additionally, its modular structure can be applied either in single- or multi-country settings, and from short to long-term macroeconomic simulations.

Efforts to extend coverage of harmonized household surveys, particularly the creation of the International Income Distribution Database (I2D2), permitted key modeling innovations in the GIDD framework. These innovations include: i) new method of recalibrating surveys’ sampling weights, ii) distinction between different sources of monetary income, iii) distinction of labor based on occupation, and iv) the use of country-specific Engel’s curves to account for changes in relative prices of food versus non-food consumption bundles.

The objective of this paper is to reintroduce the GIDD modeling framework and illustrate how it works with a single-country example. This paper is hence organized as follows. In the following section, the preliminary setting of the macro-micro linkage is described. The GIDD theoretical framework is detailed in section 2.2. More specifically, the microsimulation is implemented in five additive modules. The first module deals with changes in the demographic and education structure. The second module allows for sectoral reallocation of labor; while the third adjust relative labor incomes for different types of workers. Distribution-neutral growth is applied in module 4. Lastly, a counterfactual real household welfare aggregate is constructed by adjusting for changes in relative consumption prices.

Section 0 cover an empirical implementation for the case of Turkey. Simulations analyze the distributional consequences of increase in female labor force participation. The analysis shows `||.Stata®` routines and household survey data to recreate this modeling example will be posted on the World Banks’ GIDD website.

2 Setting the macro-micro linkage

One of the most important objectives of the GIDD framework is keeping, as consistent as possible, the evolution of common variables in the macro and microeconomic models. Large discrepancies in data and definitions between macro and micro might, nonetheless, exist and impose some natural limitations for achieving this objective. As a result, an important amount of effort is devoted to set up the inner interactions between the macro and micro models.

2.1 Linking Aggregate Variables

On top-down modelling frameworks, the macro generally leads while the micro mimics the aggregate behavior observed top. The two models must link through a set of Linkage Aggregate Variables (LAVs). In the GIDD, the set of macro-micro LAVs can be grouped in the following 7 dimensions:

1. **Geographical aggregation:** $c = \{\text{individual country or regional aggregation}\}$
2. **Time:** $t = \{0, 1, \dots, T\}$
3. **Demographics:** in a $\{m_{ct} \cdot g_{ct} \cdot s_{ct}\}$ structure with
 - $m_{ct} = \{\text{age groups}\}$
 - $g_{ct} = \{\text{gender}\}$
 - $s_{ct} = \{\text{levels of education based on completed years of schooling}\}$
4. **Labor force status:** $f = \{\text{labor force participation status}\}$
5. **Employment:** labor supply l_{pqct} , labor incomes w_{pqct} , and non-labor incomes z_{pqct} in an economy with:
 - $p = \{\text{sectors}\}$; and
 - $q = \{\text{types of workers}\}$
6. **Welfare aggregate:** aggregate income/consumption per capita $\bar{l}_{ct}$
7. **Price Index:** $\bar{p}_{bct}$ where
 - $b = \{\text{household consumption aggregates}\}$

For simplicity, we will omit the geographical dimension c for the rest of the paper. Table 1 below defines the set of LAVs employed in a typical GIDD application. Items listed in the third column from the left to right, nomenclature, will be used in Section 3 while defining the theoretical model. The last column to the right, micro application, list values used for the empirical application showcased in Section 0. In the GIDD, dynamics are driven strongly by changes in the demographic structure $\{m_t, g_t, s_t\}$. For most applications, it has been assumed that demographic changes are exogenously determined in both, the macro and micro models.

Table 2 describe a basic classification of individuals for the household survey. Each individual can be assigned to one of those mutually exclusive groups. In terms of labor force participation status, notice that the unemployed and individuals not in the labor force can conveniently be assigned to the *null* economic sector p_0 ; while employed individuals can be placed either in the shrinking, p_1 , or expanding sector, p_2 . Lastly, notice that labor is defined as more qualified and less qualified. Macro databases used for CGE models typically distinguish between these two types of labor based on occupation and we follow that convention.

Table 1 Basic LAV structure of the GIDD model

Item	Definition	Nomenclature	Micro application
t	{initial, counterfactuals}	$\{t_0, t_1\}$	{2010, 2025}
m_t	{8 age groups}	$\{m_{1t}, m_{2t}, \dots, m_{8t}\}$	{0-9, 10-19, ..., 50-59, 60+}
g_t	{gender}	$\{g_{1t}, g_{2t}\}$	{female, male}
s_t	{3 education levels}	$\{s_{1t}, s_{2t}, s_{3t}\}$	{0-6, 7-12, 13+}
f	{2 labor status}	$\{f_0, f_1\}$	{unemployed/not in labor force, employed}
p	{3 sectors}	$\{p_0, p_1, p_2\}$	{null, shrinking, expanding}
q	{2 types of workers}	$\{q_1, q_2\}$	{less qualified, more qualified}
b	{2 bundle of goods}	$\{b_1, b_2\}$	{food, non-food}

Source: Author's own elaboration

Table 2 GIDD: Basic classification of all individuals in the household survey

by labor status f	by sector of employment p	by qualification q	with labor and non-labor incomes
Unemployed / Not in the labor force	None	Less qualified	$z_{it_0p_0q_1} = 0;$ $w_{it_0p_0q_1} \geq 0$
		More qualified	$z_{it_0p_0q_2} = 0;$ $w_{it_0p_0q_2} \geq 0$
Employed	Shrinking sector	Less qualified	$z_{it_0p_1q_0} \geq 0;$ $w_{it_0p_1q_0} \geq 0;$
		More qualified	$z_{it_0p_1q_1} \geq 0;$ $w_{it_0p_1q_1} \geq 0$
	Expanding sector	Less qualified	$z_{it_0p_2q_0} \geq 0;$ $w_{it_0p_1q_0} \geq 0$
		More qualified	$z_{it_0p_2q_2} \geq 0;$ $w_{it_0p_2q_2} \geq 0$

Note: $\forall i \in h$

Source: Author's own elaboration

An important distinction is that employed individuals can receive both labor and non-labor income. The unemployed and all individuals not considered part of the labor force can only be recipients of non-labor income. In other words, the unemployed and individuals not in the labor force must have labor incomes equal to zero.

Typically, household surveys would also permit to classify employed individuals, according to their main job, as unpaid workers, paid workers, the self-employed, and employers. For the case of the self-employed and employers, labor income would ideally represent the marginal contribution from labor to total output. Capital gains, in this respect, should ideally be considered part of non-labor income. Empirically,

Table 3 Definition of variables in the macro scenario and its correspondence with micro aggregates

Linkage Aggregate Variable	Definition	Equivalence on the household survey
l_p1_q0	Labor supply in shrinking sector of less qualified	$\sum_i \theta_i \quad \forall i \in p_1 \cap q_0$
l_p1_q1	Labor supply in shrinking sector of more qualified	$\sum_i \theta_i \quad \forall i \in p_1 \cap q_1$
l_p2_q0	Labor supply in expanding sector of less qualified	$\sum_i \theta_i \quad \forall i \in p_2 \cap q_0$
l_p2_q1	Labor supply in expanding sector of more qualified	$\sum_i \theta_i \quad \forall i \in p_2 \cap q_1$
w_p1_q0	Wage index in shrinking sector for less qualified	$\frac{\sum_i \theta_i w_i}{\sum_i \theta_i} \quad \forall i \in p_1 \cap q_0$
w_p1_q1	Wage index in shrinking sector for more qualified	$\frac{\sum_i \theta_i w_i}{\sum_i \theta_i} \quad \forall i \in p_1 \cap q_1$
w_p2_q0	Wage index in expanding sector for less qualified	$\frac{\sum_i \theta_i w_i}{\sum_i \theta_i} \quad \forall i \in p_2 \cap q_0$
w_p2_q1	Wage index in expanding sector for more qualified	$\frac{\sum_i \theta_i w_i}{\sum_i \theta_i} \quad \forall i \in p_2 \cap q_1$
I	Per capita household income aggregate	$\frac{\sum_i \theta_i y_i}{\sum_i \theta_i}$
P_b1	Price index for food	1
P_b2	Price index for non-food	1

Source: Authors' own elaboration

nonetheless, quality and scope of harmonized survey data usually limits the analysis and force the microsimulation to rely on a set of assumptions. Section 3 will expand on these aspects.

2.2 Macro scenario files

Scenario files contain a set of time-varying macroeconomic variables that signal country-specific changes. A typical scenario file contains one prospective long-term macroeconomic trajectory and hence portrays a path of economic development that follows some previously assumed premises. As macro modelers know well, when one of such premises varies, a new scenario is generated. As a result, there are infinite number of possible macro scenarios and while some might produce drastically different socioeconomic pathways, others might result negligible or irrelevant.

Macro modelers interested in sending responses to the GIDD should prepare a file that contain, for each country c to be simulated, the set of linkage aggregate

variable shown in Table 3 above. The macro scenario file used for simulation in Section 4 is available for download on the World Bank's GIDD website.

3 Macro-micro simulation steps

The macro-micro simulation framework has 5 broad steps explained through this section. The first step changes the micro data demographic structure. The second step allows labor to migrate from a shrinking to an expanding economic sector. Changes in sectoral wage premia are modelled in the third step. In step four, distribution-neutral growth is applied. Lastly, changes in relative consumption prices reflect changes in real incomes. Let's start by formally defining nominal income of individual i in household h as the sum of labor, w_{iht} and non-labor, z_{iht} income

$$y_{thi} = w_{thi} + z_{thi} \quad \forall i \in h \wedge t = \{t_0, t_1\} \quad (1)$$

As a result, household total nominal income follows from Equation (2)

$$Y_{ht} = \sum_{i \in h} y_{iht} \quad (2)$$

Real household total income is obtained dividing the previous expression by a price deflator $P_{ht}(\alpha_h, \bar{P}_{b_0t}, \bar{P}_{b_1t})$:

$$Y_{ht}^P = \frac{Y_{ht}}{P_{ht}(\alpha_h, \bar{P}_{b_0t}, \bar{P}_{b_1t})} = \frac{Y_{ht}}{\alpha_{ht}\bar{P}_{b_0t} + (1 - \alpha_{ht})\bar{P}_{b_1t}} \quad (3)$$

where α_{ht} is the budget share spent on food by household h . Typically, a macro model indexes initial consumer prices equal to 1 so that:

$$\bar{P}_{b_0t_0} = \bar{P}_{b_1t_0} = 1 \quad (4)$$

which results in

$$Y_{ht_0}^P = Y_{ht_0} \quad (5)$$

In the household survey, total household per capita income $\bar{I}_{t_0}$ can be obtained from Equation :

$$\bar{I}_{t_0} = \frac{\sum_h \theta_{ht_0} Y_{ht_0}^P}{\sum_h \theta_{ht_0}} \quad (6)$$

where θ_{ht_0} is the household-specific frequency sampling weight.

3.1 Changes in the demographic structure

The first step in the microsimulation is to establish a common set of pathways for the demographic structure $\{m_t \cdot g_t \cdot s_t\}$ in both, the macro and micro models. These counterfactual demographic pathways are exogenously determined. Note that $\{m_{t_0} \cdot g_{t_0} \cdot s_{t_0}\}$ can be obtained directly from the household survey, while

counterfactual demographic structures for t_0 to t_N can be constructed relying on external sources. The application showcased in Section 4, for instance, builds counterfactual demographic pathways using the Shared Socioeconomic Pathways (SSP) database. The SSP database provides population projections by age, education, and education².

Once the counterfactual demographic structure $\{m_{t_1} \cdot g_{t_1} \cdot s_{t_1}\}$ has been set, the objective is to obtain a set of household sampling weights that replicate the counterfactual demographic structure. The GIDD relies on the minimum-cross entropy method proposed by Wittenberg (2010). This method allows to recreate homogenous weights within household members. Consider that in most household surveys, the unit of analysis and sampling is the household and hence sampling weights attached to every individual within the household are equal. It follows that:

$$\sum_h \sum_i \theta_{iht_0} = H_{t_0} \quad (7)$$

in which H_{t_0} is the total population in time t_0 and

$$\theta_{iht_0} = \theta_{jht_0} = \theta_{ht_0} \quad \text{for } i \neq j \wedge i, j \in h$$

where θ_{iht_0} represents the initial frequency weight of individual i within household h , equal to the common weight θ_{ht_0} . So, equation (7) can be rewritten as equation (8):

$$\sum_h \theta_{ht_0} \cdot \text{household size}_h = \sum_h \theta_{ht_0}^* = H_{t_0} \quad (8)$$

Equation (8) can be written in terms of probabilities:

$$1 = \sum_h \frac{\theta_{ht_0}^*}{H_{t_0}} = \sum_h \vartheta_{ht_0}^* \quad (9)$$

Given a new structure of total population to estimate:

$$\sum_h \sum_i \tilde{\theta}_{iht_1} 1[m_{t_1}, g_{t_1}, s_{t_1}] = \tilde{H}_{t_1} 1[m_{t_1}, g_{t_1}, s_{t_1}] \quad (10)$$

in which $1[m_{t_1}, g_{t_1}, s_{t_1}]$ is the indicator function, $\tilde{H}_{t_1}$ is the new total population, and each member i of household h has the same weight,

$$\tilde{\theta}_{iht_1} = \tilde{\theta}_{jht_1} = \tilde{\theta}_{ht_1} \quad \text{for } i \neq j \wedge i, j \in h \quad \text{then}$$

² It is important to highlight that there might exist discrepancies between the $\{m_{t_0} \cdot g_{t_0} \cdot s_{t_0}\}$ structure found in the microdata and any other initial structure $\{\hat{m}_{t_0} \cdot \hat{g}_{t_0} \cdot \hat{s}_{t_0}\}$ collected from external sources.

$$\sum_h \tilde{\theta}_{ht_1} \text{household size}_h = \sum_h \tilde{\theta}_{ht_1}^* = \tilde{H}_{t_1} \quad (11)$$

Equation (12) expresses this in probability terms:

$$1 - \sum_h \frac{\tilde{\theta}_{ht_1}^*}{\tilde{H}_{t_1}} = \sum_h \tilde{\vartheta}_{ht_1}^* \quad (12)$$

The idea is to go from the initial household weight ($\theta_{ht_0}^*$) or probability ($\vartheta_{ht_0}^*$) distribution to a new one ($\tilde{\theta}_{ht_1}^*$ or $\tilde{\vartheta}_{ht_1}^*$) imposing several constraints such as education, gender, or age-group population growth rates. Formally, the new probability distribution can be obtained by solving the optimization problem in equation (13)

$$\min \Gamma(\vartheta, \tilde{\vartheta}) = \sum_h \tilde{\vartheta}_{ht_1}^* \ln \left(\frac{\tilde{\vartheta}_{ht_1}^*}{\vartheta_{ht_0}^*} \right) \quad (13)$$

such that

$$E(x_{rt}) = \sum_h \vartheta_{ht_0}^* k_{x_{rt}h} \quad r=1, \dots, R \quad (14)$$

$$\sum_h \vartheta_{ht_0}^* = 1$$

The term $k_{x_{rt}h}$ is the mean of the x_r characteristic within the household h ($k_{x_{rt}h} = \frac{\sum_i x_{ih}}{h \text{size}_h}$). In other words, this optimization procedure asserts that the distribution $\tilde{\vartheta}_{ht_1}^*$ that meets the moment constraints and the normalization restriction while requiring the least additional information should be picked, and the one that deviates as little as possible from the initial distribution ($\vartheta_{t_0}^*$) should be picked. The solution to this problem is given by equation (15)

$$\hat{\vartheta}_{ht_1} = \vartheta_{ht_0}^* \frac{\exp(\chi \hat{\lambda})}{\Omega(\hat{\lambda})} \rightarrow \sum_h \hat{\vartheta}_{ht_1} = \sum_h \theta_{ht_0}^* \varrho_h = \tilde{H}_{t_1} \quad (15)$$

$$\text{with } \varrho_h = \frac{\exp(X \hat{\lambda})}{\Omega(\hat{\lambda})}$$

in which $\hat{\lambda}$ is the Lagrange multiplier for the constraint $E(X)$, and Ω is the normalizing factor to scale the new population to the target size.

3.2 Labor supply and sectoral reallocation

Initial labor supply, l_{pqt_0} and labor income, w_{pqt_0} , can be obtained from the household level data. Nevertheless, their counterfactual levels in t_1 will be given by the macro general equilibrium model. Workers will choose to abandon the shrinking sector if this choice represents an increase in their expected earnings. The

microeconomic model faces the constraint of moving l workers out of the shrinking into the expanding sector. The GIDD selects migrants from the shrinking sector.

The probability of observing an individual i working in the expanding sector ($p=p_1$) can be modelled using a probit equation (16):

$$Pr(p = p_1) = P(X_i, Z_i) \quad (16)$$

where X_i and Z_i are vectors of individual and household characteristics for individual i , respectively. Workers in their shrinking sectors, p_0 , can be assigned a score based on their set of characteristics X and Z and sorted accordingly. Workers in the shrinking sector, p_0 , with the higher probability of being in the expanding sector, p_1 , can migrate up to a point where the predicted share of labor is satisfied. Let's assume that less qualified labor, q_1 , is allowed to migrate, then equation below can describe the from the macro model.

$$\delta = \frac{\delta_{t_1}}{\delta_{t_0}} = \frac{\frac{l_{p_1 q_0 t_1}}{l_{p_1 q_0 t_1} + l_{p_1 q_0 t_1}}}{\frac{l_{p_1 q_0 t_0}}{l_{p_1 q_0 t_0} + l_{p_1 q_0 t_0}}} \quad (17)$$

where $\delta > 1$. Similarly, on the micro side we can define the labor supply of workers as:

$$\begin{aligned} l'_{pqt_0} &= \sum_h \sum_{i \in pqh} \theta_{ih} \\ l'_{pqt_1} &= \sum_h \sum_{i \in pqh} \tilde{\theta}_{ih} \end{aligned} \quad (18)$$

where

$$\delta' = \frac{\delta'_{t_1}}{\delta'_{t_0}} = \frac{\frac{l'_{p_1 q_0 t_1}}{l'_{p_1 q_0 t_1} + l'_{p_1 q_1 t_1}}}{\frac{l'_{p_1 q_0 t_0}}{l'_{p_1 q_0 t_0} + l'_{p_1 q_1 t_0}}} \quad (19)$$

If $1 \leq \delta' < \delta$, the total number of workers that exist in the expanding sector can be found:

$$l^* = \delta'_{t_0} \frac{\delta t_1}{\delta t_0} (l'_{p_1 q_0 t_1} + l'_{p_1 q_1 t_1}) - l'_{p_1 q_0 t_1} \quad (20)$$

By allowing the new demographic structure to operate. If $l^* > 1$, GIDD assumes that there is an extra amount of labor that can migrate. Nevertheless, when $l^* < 1$, indicates that reweight is making as a side-effect, a larger sectoral reallocation than the CGE.

3.3 Changes in sectoral wage premia

The third step is to adjust factor returns by skill and sector in accordance with the results of the CGE model. The GIDD imposes an entirely new vector of earnings on each worker, conditional on that worker being in sector p and qualification level q . In order to ensure consistency between the macro and micro outcomes, the GIDD adjusts the ratios between wage premia rather than wages themselves.

Beginning with a distribution of earnings from labor, w , by sector, p , and qualification, q , in the macro data, define a set of wage ratios as follows:

$$g_{pqt_0} = \frac{w_{pqt_0}}{w_{p_0q_0t_0}} \quad (21)$$

where $w_{p_0q_0t_0}$ are the initial average earnings from labor of the numeraire for the less qualified workers in the shrinking sector. Similarly, a set of wage ratios for the macroeconomic counterfactual scenario can be defined as follows:

$$\hat{g}_{pqt_1} = \frac{w_{pqt_1}}{w_{p_0q_0t_0}} \quad (22)$$

where w_{pqt_1} and $w_{p_0q_0t_1}$ represent wage values from the CGE model in the macroeconomic counterfactual scenario. In general, the micro data will have a set of initial wage ratios g'_{pqt_0} that will differ from the macro data's g_{pqt_0} . Analogous to equation 5, define:

$$g'_{pqt_0} = \frac{\sum_h (w'_{pqt_0} \cdot \vartheta_{ht_0}^*)}{\sum_h (w'_{p_0q_0t_0} \cdot \vartheta_{ht_0}^*)} \quad (23)$$

where $g'_{s,e}$ represent weighted wage ratio by qualification, q , and sector, p ; w'_{pqt_0} are labor earnings obtained from household survey data. Similarly

$$g'_{pqt_1} = \frac{\sum_h (w'_{pqt_1} \cdot \tilde{\vartheta}_{ht_1}^*)}{\sum_h (w'_{p_0q_0t_1} \cdot \varphi_{ht_1}^*)} \quad (24)$$

The counterfactual wage ratios $\hat{g}'_{s,e}$ for the microdata will then be calculated as:

$$\hat{g}'_{pqt_1} = g'_{pqt_0} \frac{\hat{g}_{pqt_1}}{g_{pqt_0}} \quad (25)$$

This implies that even if initial and final wages differ between the macro and micro models, the change in wage ratios will be consistent across the two models. This diminish the possibility of wage gap reversal and ensures that the distributional changes are consistently mapped from the macro to the micro data.

3.4 Economic growth

Note that Equation 10 only operates on labor income. In order to adjust the micro data such that the weighted average percentage change in the per capita income/consumption across all households matches the change in real consumption per capita in the CGE model, a subsequent adjustment is carried out. Define Y as real per capita income calculated from the CGE model in the benchmark and $\hat{Y}$ as

its predicted value in the CGE model simulation. Define $\gamma'_h = \sum_i y'_{i,h} / n_h$ as the per capita income of household h , where i identifies each household member, and n is equal to the size of household h . Then let's define Y' as the weighted average value of real per capita income across all households, i.e.,

$$\sum_h v_h \gamma'_h = Y' \quad (26)$$

where v_h is the weight of household h in aggregate income. Correspondingly

$$\sum_h \omega_h \lambda \hat{\gamma}'_h = \hat{Y}' \quad (27)$$

is the weighted average per capita income value in the policy simulation. Note that $\sum_h v_h = 1$, $\sum_h \omega_h = 1$ and λ is a scalar. Equations 11 and 12 allow for different household weights since the weights of the households will typically change over time. We constrain $\hat{Y}'$ by equation 13:

$$\hat{Y}' = Y' \frac{\hat{Y}}{Y} \quad (28)$$

We implement this constraint in a distribution neutral way. That is, we adjust all household income in the counterfactual by a scalar λ such that per capita household income equals $\lambda \gamma'_h$: as a result, λ can be defined by:

$$\lambda \sum_h \omega_h \hat{\gamma}'_h = Y' \frac{\hat{Y}}{Y} \quad (29)$$

This third step in the microsimulation assures that national income grows accordingly to the CGE model. Despite the fact that the GIDD ignores other forms of income, such as capital income, this transformation guarantees consistency between the weighted average household income assessment and the CGE model assessment. For poor households, which is the main focus of the work, the assumption should be reasonably accurate since poor households have little capital income. There is more of a margin of error for wealthier households. But for these households, it is skilled labor rather than unskilled labor that tends to be more important and Bussolo, de Hoyos, Medvedev (2010) have noted a tendency for skilled wage and returns to capital to be correlated.

3.5 Relative Prices

Finally, macroeconomic estimates of changes in agricultural and non-agricultural prices are distributed across heterogeneous households using the following method. Let us define the initial per capita monetary income of household h , γ'_h , and the purchasing power of household h , γ_h^r , as the ratio of its monetary income divided by a household-specific price index capturing the household's consumption patterns in terms of food and non-food expenditure:

$$Y_h^r = \frac{\gamma_h'}{P_h} = \frac{\gamma_h'}{\alpha_h P_f + (1 - \alpha_h) P_{nf}} \quad (30)$$

where P_f and P_{nf} are food and non-food price indices and α_h is the proportion of household's h budget spent on food. Equation 14 captures the dual effect of a price variation, i.e. higher monetary income on the one hand, and the loss of purchasing power on the other.

The α_h parameter in the denominator of the right hand side of Equation 14 can be estimated with household data using the following specification:

$$\alpha_h = \beta_0 + \beta_1 \ln(\gamma_h') + e_h \quad (31)$$

where e_h is a vector of household-specific errors that are assumed to be distributed with $E(e_h) = 0$ and $V(e_h) = \sigma^2$. Assuming that estimated parameters $\hat{\beta}_0$ and $\hat{\beta}_1$ remain constant, the new budget share spent on food for household h , α_h' , at the counterfactual per capita income, $\lambda \hat{\gamma}_h'$, can be obtained from:

$$\alpha_h' = \hat{\beta}_0 + \hat{\beta}_1 \ln(\lambda \hat{\gamma}_h') + \hat{e}_h \quad (32)$$

The changes in real per capita incomes brought about by a change in relative prices of food versus non-food can be approximated by the following linear expression:

$$\hat{\gamma}_h^r = \frac{\lambda \hat{\gamma}_h'}{\alpha_h' P_f' + (1 - \alpha_h') P_{nf}'} \quad (33)$$

where $\hat{\gamma}_h^r$ in Equation 17 is the real per capita income adjusted for changes in relative prices of food versus non-food. $\hat{\gamma}_h^r$ is the counterfactual welfare aggregate for the analysis of poverty and shared prosperity.

4 An application for the case of Turkey

This section describes the GIDD simulation results for the case of Turkey.

- Baseline assumes unprecedented favorable conditions in terms of household income growth.
- The benefits of growth can be observed in the complete eradication of extreme poverty.
- As a result, the pure growth effect dominates over other distributional effects resulted from policy changes.
- Even in this conditions, it is possible to decompose the distributional consequences of each scenario.

First, data sources and definitions of key variables are discussed. Section 4.2 covers the distributional consequences of each simulations step.

4.1 Basic data and definitions

4.1.1 Macroeconomic scenarios

Macro scenarios were obtained from a recursive-dynamic general equilibrium model for the Turkish economy. Macro scenarios provide the set of linkage aggregate variables needed for the microsimulation, including population and skill formation, which are exogenously produced by combining population totals from UN World Population Prospects and skill formation from scenario 1 of the Shared Socio Economic Pathways Database. The general equilibrium model is a variant from Griffin () for the case of Poland. A baseline and 3 alternative scenarios are modelled for the period 2010–2025. It is assumed that baseline and alternative scenarios are identical from 2010 to 2015. Policy changes start in 2016.

0. Baseline conditions
1. Increase in capital tax rate & household savings rates
2. Increase in current account balance & 2% increase in TFP rates
3. Increase of 1.5% on total factor productivity for more skilled labor

4.1.2 Microdata

The microsimulations are performed on top of Turkey’s Household Budget Survey (HBS) for the year 2010³. The survey has been obtained from the World Bank microdata catalog and was previously harmonized by World Bank staff.

This exercise requires to differentiate labor according to skill level and formality. Using definitions from HBS, labor can be classified as follows.

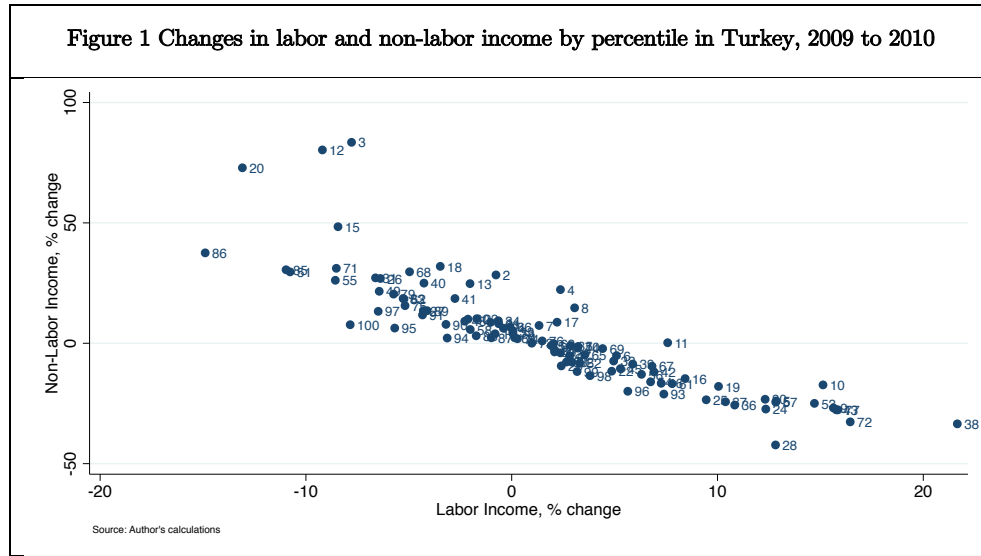
³ Official information for this survey can be obtained from Turkish Statistical Institute website http://www.turkstat.gov.tr/Kitap.do?metod=KitapDetay&KT_ID=7&KITAP_ID=256. Prospective users of World Bank’s harmonized microdata must ask for access to dataliweb@worldbank.org

- **Less skilled labor:** Not literate, Literate and did not finish school, Primary school, Primary, Middle school, Middle school, Middle vocational, High school, Lise vocational.
- **More skilled labor:** individuals with 2-year colleges, 4-year colleges faculty, Master and PhD.
- **Informal:**
- **Formal:**

4.1.3 Modelling Non-labor income

The treatment of labor and non-labor income is not a trivial modeling exercise. In previous GIDD applications and mainly to data constraints, non-labor income was assumed to grow at the same rate as labor-income. In this application, nonetheless, we follow a different approach. Instead of allowing non-labor to grow at the same rate as labor income, gains in labor income would decline the share of non-labor in total household income. In other words, households tend to smooth consumption by complementing between labor and non-labor income.

Figure 1 show changes in labor and non-labor income at the percentile level derived from repeated cross-sections using HBS for 2009 and 2010.



4.1.4 Initial data discrepancies

Reconciliation of macro and micro data is arguably the most difficult task during macro-micro simulations. Initial discrepancies in key linkage aggregate variables are common. In most cases they emerge as the natural consequence of gathering data from complimentary sources. In our case, the macro model collected labor data from the 2010 Turkish Labor Force Survey (LFS). Underlying differences in design and data collection yield to large discrepancies between the LFS and HBS surveys, particularly in terms of labor force participation rates, skill level of the labor force, and prevalence of informality. The HBS will follow paths marked by the macroeconomic model.

4.2 Poverty and Shared Prosperity

This section explores the simulation results in terms of progress on reducing poverty and achieving shared prosperity.

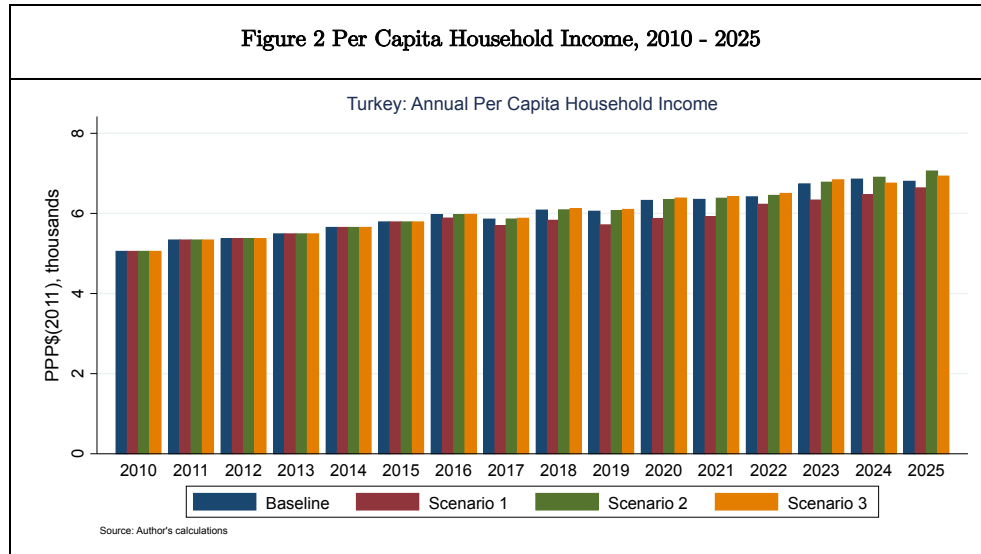
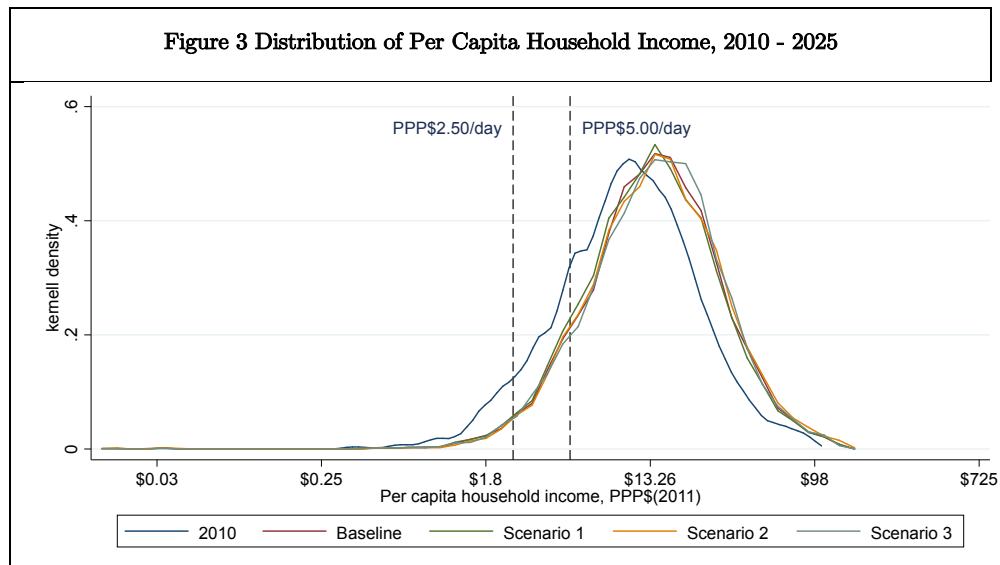
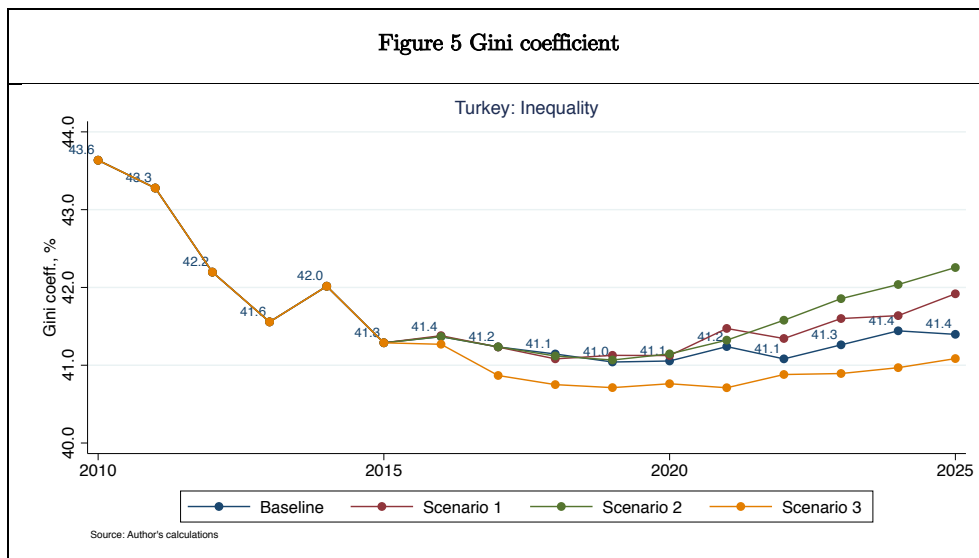
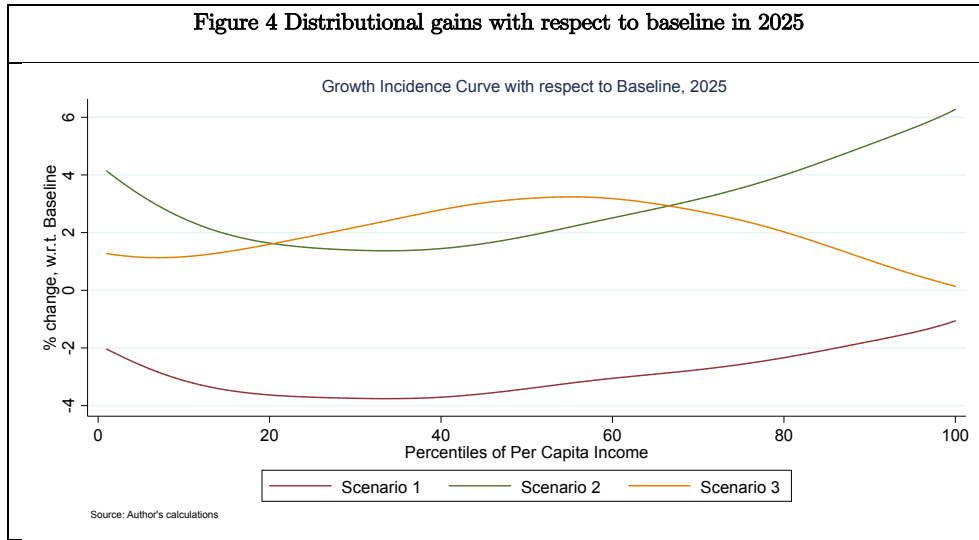
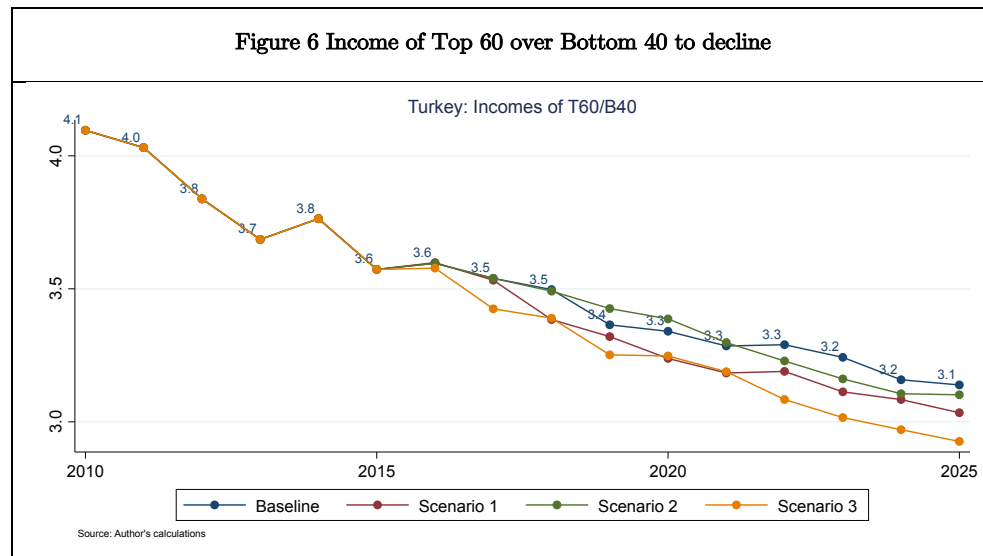


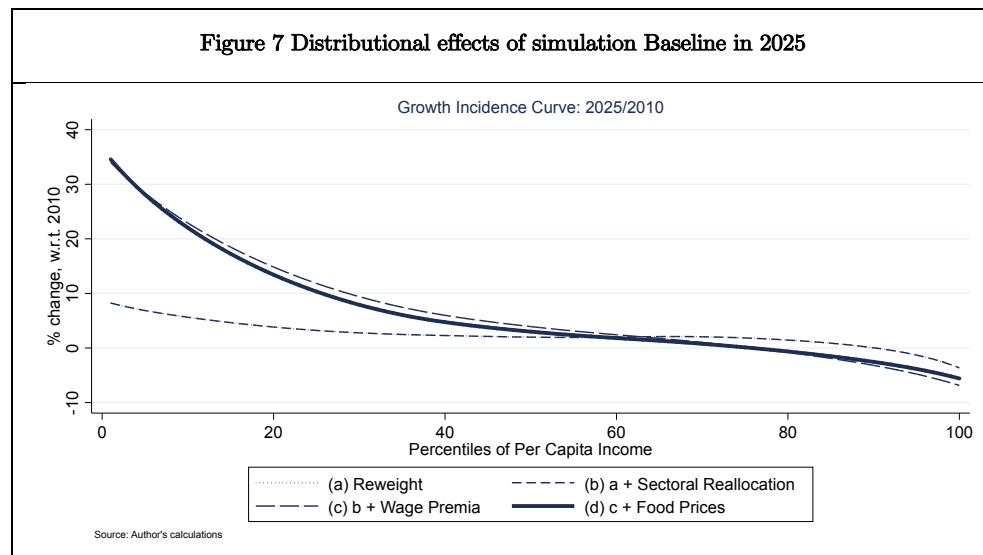
Figure 1 shows average household income per capita. For 2010, harmonized microdata from World Bank reports an average household income of \$5,050 in PPP(2011) international dollars, which will be taken as the starting point in our simulations. Baseline and alternative scenarios are assumed to be identical from 2010 to 2015. Average household income per capita grows rapidly at a 10.17% annual rate. Starting in 2016, the macro model assumes an annual growth rate of 8.89% in per capita household income for the baseline scenario.





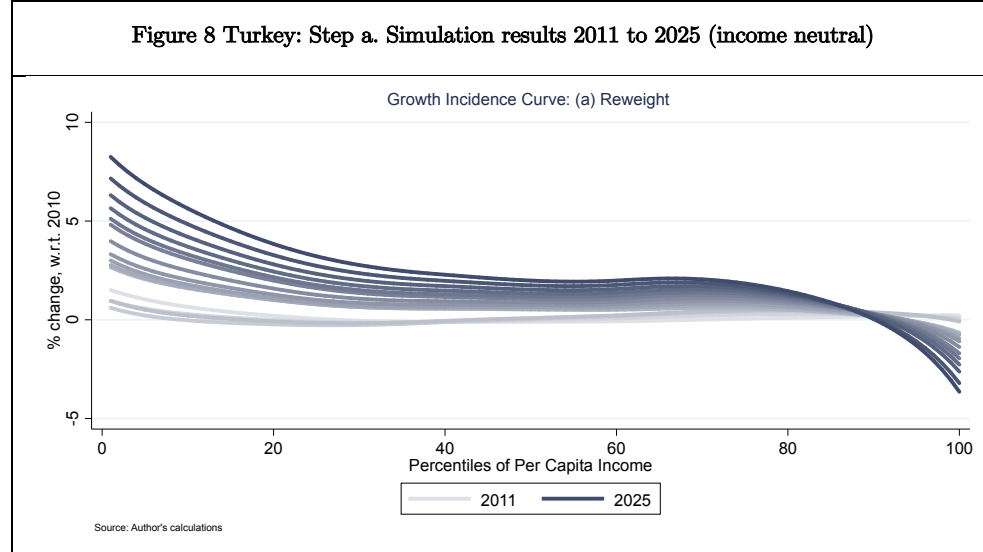


4.3 Decomposing simulation steps



As the population profile changes, particularly as it ages and reaches higher levels of education, the process of recalibrating the sampling weights creates *side-effects* on all other variables in the household survey. Let's suppose that informality is concentrated on the younger and less-educated type of workers; if we aimed to match the population profile of a future target year that anticipates population ageing and higher levels of education, then recalibrating the sampling weights will naturally

reduce the rate of informality⁴. The figure below depicts growth incidence curves that show the effect of the reweight process on household per capita income across time.

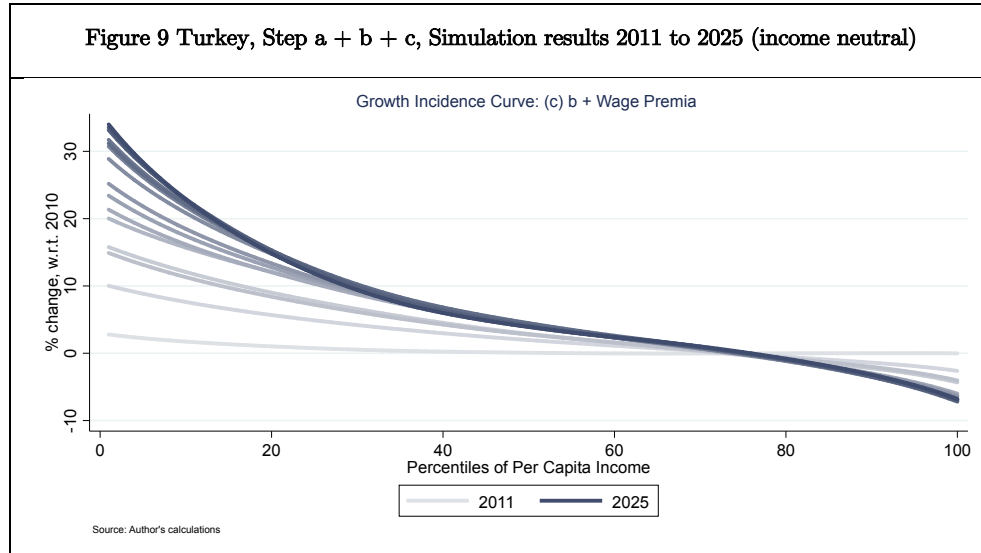


Sectoral reallocation allows labor to migrate from the shrinking (informal) to an expanding (formal) sector. Labor to be moved can be chosen either randomly or according to their individual characteristics. The amount of labor that will be moved out of informality is given exogenously by the CGE.

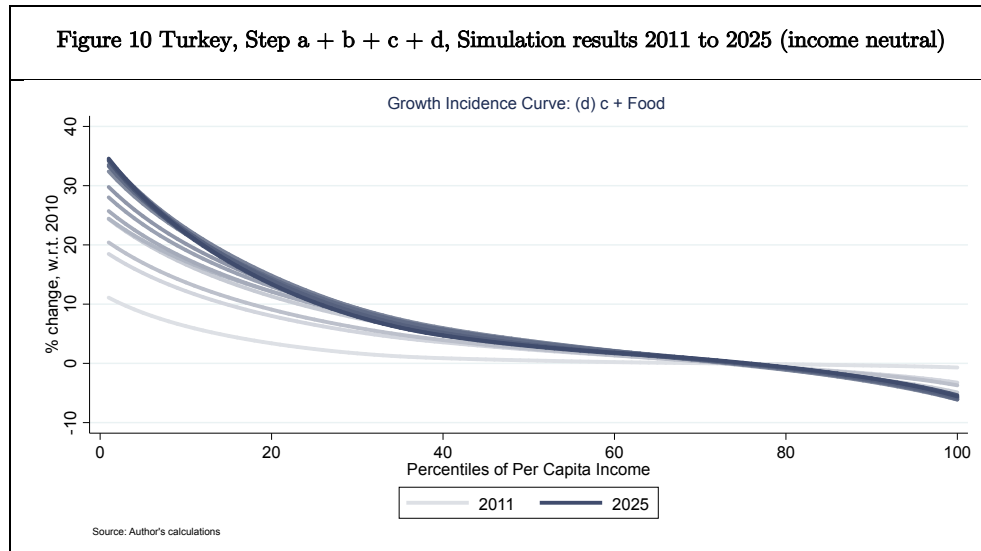
So far, the *side-effect* discussed above is creating a slightly larger reduction in the informal sector than the one projected by the CGE. The microsimulation model prevents reverse labor mobility. I will revise more carefully this step. As of now, this step has been effectively switched-off and won't have distributional consequences.

For workers defined in a 2x2 matrix: 2 types of workers (more educated, less educated) and 2 sectors (shrinking and expanding), the microsimulation adjust labor incomes for each type of worker. In the current scenarios, the effect of changes in wage premia dominate over the effects of reweight. Over time the effect of wage premia is larger and benefits relatively workers in the lower part of the income distribution. Figure 2 below shows the cumulative effects of Steps a + b + c.

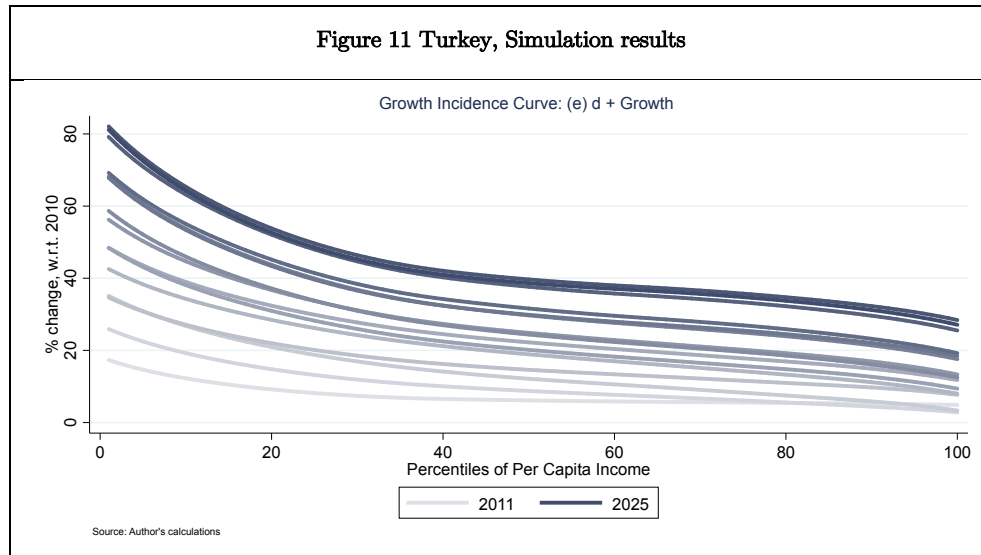
⁴ Additionally, there exist an initial side effect attributable to differences between macro targets and microdata.



Change in food prices is applied on top of the household per capita income and operates as a real price deflator. Despite the fact that the CGE provides higher relative prices of food versus non food, this regressive distributional effect is small in comparison with changes in wage premia. Figure 3 below show effect of Steps a + b + c + d



Household Consumption Growth. The final step is simply a distributional neutral shift for all households according to household consumption growth given by the CGE. In these scenarios, growth is largely the most important shock on the baseline scenario. Figure 3 below shows the final growth incidence curve that contains all simulation results, from Step a + b + c + d + e



5 Conclusions

This paper presented an expanded version of the GIDD model. New household surveys permitted innovations in the modeling framework. Further improvements and open access to data and underlying code make these modeling tools widely available. Simulation results used for the case of Turkey show [To be incorporated].

Bibliography

- Ahmed, Amer S., Marcio Cruz, Delfin S. Go, Maryla Maliszewska, and Israel Osorio Rodarte. 2014. "How Significant Is Africa's Demographic Dividend for Its Future Growth and Poverty Reduction?"
- Balistreti, Edward J., Maryla Maliszewska, Israel Osorio Rodarte, David G. Tarr, and Hidemichi Yonezawa. 2016. "Poverty and Shared Prosperity of Deep Integration in Eastern and Southern Africa." Policy Research Working Paper Series.
- Bourguignon, François, and Maurizio Bussolo. 2013. "Income Distribution in Computable General Equilibrium Modeling." In , 1:1383–1437. Handbook of Computable General Equilibrium Modeling. Elsevier. <http://ideas.repec.org/h/eee/hacchp/v1y2013icp1383-1437.html>.
- Bourguignon, Francois, Maurizio Bussolo, and Luiz A Pereira da Silva. 2008. *The Impact of Macroeconomic Policies on Poverty and Income Distribution: Macro-Micro Evaluation Techniques and Tools*. Houndmills, Basingstoke, Hampshire ; New York Washington, DC: Palgrave Macmillan ; World Bank. <http://www.loc.gov/catdir/toc/ecip081/2007040478.html>.
- Bourguignon, François, Francisco H. G. Ferreira, and Phillippe G. Leite. 2008. "Beyond Oaxaca–Blinder: Accounting for Differences in Household Income Distributions." *The Journal of Economic Inequality* 6 (2): 117–48. doi:10.1007/s10888-007-9063-y.
- Bussolo, Maurizio, Rafael E De Hoyos, and Denis Medvedev. 2010. "Economic Growth and Income Distribution: Linking Macro-Economic Models with Household Survey Data at the Global Level." *International Journal of Microsimulation* 3 (1): 92–103. <http://ideas.repec.org/a/ijm/journal/v3y2010i1p92-103.html>.
- Davies, James B. 2009. "Combining Microsimulation with CGE and Macro Modelling for Distributional Analysis in Developing and Transition Countries." *International Journal of Microsimulation* 2 (1): 49–65.
- Devarajan, Shantayanan, Delfin S. Go, Maryla Maliszewska, Israel Osorio-Rodarte, and Hans Timmer. 2015. "Stress-Testing Africa's Recent Growth and Poverty Performance." *Journal of Policy Modeling* 37 (4): 521–47. doi:10.1016/j.jpolmod.2015.04.006.
- van der Mensbrugghe, Dominique. 2011. "LINKAGE Technical Reference Document Version 7.1." http://siteresources.worldbank.org/INTPROSPECTS/Resources/334934-1314986341738/TechRef7.1_01Mar2011.pdf.
- Wittenberg, Martin. 2010. "An Introduction to Maximum Entropy and Minimum Cross-Entropy Estimation Using Sta." *The Stata Journal* 10 (3): 315–30.