Estimating the Value Added Tax Gap in Tanzania

An Empirical Analysis

Amina Ebrahim¹, Ezekiel Swema², Oswald Haule², Vincent Leyaro³, and Sebastián Castillo⁴

1. UNU-WIDER; 2. Tanzania Revenue Authority; 3. University of Dar es Salaam; 4. University of Helsinki and FIT
Summary

By gathering administrative tax and auditing data, we estimate the VAT gap for 2014-2019.

1. We study audited firm data to provide a preliminary evidence.
   - We document evidence about strategic behavior to avoid stronger audit prob. and evade more.

2. We follow a bottom-up approach.
   - Our preferred estimation shows a VAT gap of 48% for the audited tax region and 53% for the country.
   - The VAT gap is decreasing over time but becomes stable for the period 2016-2019.
   - This is a lower-bound estimation.

3. Using the evasion prediction, we study the behavior of firms.
   - Firms at the beginning of VAT distribution evade more.
   - Small and large-sized firms evade more.
   - We document evidence that firms declare more VAT purchases to increase evasion. This produces smaller VAT declarations, mimicking small-sized firms.
Institutional Background

• VAT is collected by the Tanzania Revenue Authority (TRA).
• Tanzania is divided into geographical localities (cities) and tax regions -> One city could have more than one tax region and vice-versa.
• Indirect tax rates are:
  • Zero-rated -> Declared in VAT act 2014. Include some exports and imports.
  • Exempted -> Declared in VAT act 2014.
  • VAT -> 18% in Mainland Tanzania.

• VAT:
  • Annual gross sale > TZS 40 million to be registered as a VAT agent.
  • Turnover > TZS 100 million must be registered for VAT -> Mandatory to have a VAT ID.
  • Tanzania has a credit-invoice computation method -> VAT = sales – purchases.

• Audits:
  • Annual audit plan by each tax region.
  • Auditing is based on taxpayer turnover trends and payments -> Risk assessment estimation.
Data

  - Date of auditing and period covered.
  - Type of auditing.
  - Amount recovered by type of tax -> We can identify “VAT compliers”.

- VAT declarations (sales and purchases) at the firm level between 2011-2021.
  - Monthly declarations in the VAT form.
  - Taxed and untaxed (exemptions and zero-rated) items.
  - Gross and tax-paid amounts. -> For some items.

- Firms’ information.
  - Firm’s ID.
  - Firm’s VAT ID.
  - Tax region, postal city, business activity, and industry (ISCI 4-digit code).
Data

- Audited tax regions have more firms, but the rate of audited firms is around 15%. The audited tax regions show a larger output, inputs, and VAT payment rate.
- Firms in audited and unaudited tax regions **bunch around zero** VAT declaration. VAT recovered is significant compared with the auditing rate (average 15%).
Preliminary Evidence

VAT evasion

- Firms with positive and negative VAT declarations and large-sized firms show more auditing processes.
- Evasion is not so different regarding VAT declarations, but firms that bunch at zero evade more. For sales, this is similar, and small and large-sized firms evade more.
- Using the rate evasion/sales, firms that bunch at zero VAT declaration and small-sized firms evade more.
Empirical Strategy

• For audited firms, we estimate

\[ \text{Evasion}_{ikt} = \alpha + \beta_1 O' + \beta_2 I' + \beta_3 \text{Net Profits}_{ikt} + \lambda_1 \text{Vat}_{ikt} + \lambda_2 \text{Sales}_{ikt} + \lambda_3 date_t + \lambda_4 \text{Tax region}_{ik} + \lambda_5 \text{Activity}_i + \lambda_6 \text{City}_{ikt} + \lambda_7 \text{ISIC4}_{ikt} + u_{ikt} \]

where O is the sales inputs from the VAT form, and I is the purchase inputs. VAT and Sales are variables to capture the distribution of both variables.

• Using the estimated coefficients, we predict evasion in non-audited firms.

• We estimate the tax gap in the following way

\[ \text{VAT GAP} = \frac{\sum \text{Tax Evasion}}{\sum \text{Tax Evasion} + \sum \text{Tax Paid}} \]

where evasion can be the discovered amount and the estimation.

• We sum each variable per year.

• To estimate the country VAT gap, we assume the tax declaration rate between audited tax regions and country (audited/country) is the same for the VAT gap.

• To avoid calculation problems, we use absolute values in the tax declaration.
Results
Audited tax regions

• Preferred estimation, using only positive evasion predictions.
• The average VAT gap decreases across the years.
• Average VAT gap: 48.5%.
• Between 2016-2019 became more stable, reaching an average of 44%.
• Without LTD, group estimations rise 20pp approx.
• This shows we are obtaining a lower-bound.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VAT due</td>
<td>24.99</td>
<td>17.95</td>
<td>16.61</td>
<td>11.1</td>
<td>7.72</td>
<td>9.59</td>
<td>14.66</td>
</tr>
<tr>
<td>VAT</td>
<td>60.57</td>
<td>53.86</td>
<td>48.76</td>
<td>44.58</td>
<td>38.12</td>
<td>45.26</td>
<td>48.53</td>
</tr>
<tr>
<td>VAT proxy</td>
<td>63.67</td>
<td>50.00</td>
<td>48.29</td>
<td>43.77</td>
<td>37.84</td>
<td>42.71</td>
<td>47.71</td>
</tr>
</tbody>
</table>

VAT due, VAT, VAT proxy, Output VAT
Results

Country

- Same patterns as before.
- Average VAT gap: 53%.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VAT due</td>
<td>28,29</td>
<td>19,67</td>
<td>18,35</td>
<td>12,35</td>
<td>8,37</td>
<td>10,79</td>
<td>16,30</td>
</tr>
<tr>
<td>VAT</td>
<td>63,11</td>
<td>57,62</td>
<td>54,32</td>
<td>48,92</td>
<td>44,82</td>
<td>49,37</td>
<td>53,03</td>
</tr>
<tr>
<td>VAT proxy</td>
<td>66,02</td>
<td>53,35</td>
<td>52,62</td>
<td>48,93</td>
<td>41,02</td>
<td>51,24</td>
<td>52,20</td>
</tr>
<tr>
<td>Output VAT</td>
<td>22,95</td>
<td>17,86</td>
<td>18,98</td>
<td>16,24</td>
<td>14,92</td>
<td>15,32</td>
<td>17,71</td>
</tr>
</tbody>
</table>
Results
Firms’ Behavior

Normalizing by sales, firms evade more at the beginning of VAT declaration distribution. Regarding sales, evasion has a U-shape.
Preliminary conclusions

Results:

• Our preferred estimation shows a VAT gap of 48% for the audited tax region and 53% for the country.

• This is a lower-bound estimation.

• We document evidence that firms declare more purchases VAT to increase evasion. This produces smaller VAT declarations and mimics small-sized firms.

Policy Implications:

• VAT gap prediction needs to consider heterogeneity across VAT declaration and sales.

• Evasion is monotonous across VAT declaration but not across sales. Large-sized firms need attention.

• Firms are strategic agents, and auditing needs to be perceived as random as possible.
Next Steps.....

• Improve estimation of evasion:
  • Machine Learning to improve evasion predictions.
  • Model á-la-Heckman: Implicit function that determines evasion.
  • Censored model: Tobit, other? -> Evasion is censored: $y_i \geq 0$

• Study the determinants of evasion:
  • Event study design.
  • Staggered DiD.
  • Study audit and compliance probability.

• Estimate the revenue consequences of evasion.
Thank you

Sebastián Castillo Ramos

Any comments are welcome, reach me at sebastian.castillo@helsinki.fi
Data

Firms in audited and unaudited tax regions bunch around zero VAT declaration. VAT recovered is significant compared with the auditing rate (average 15%).
Preliminary Evidence
Number of firms

Firms with positive and negative VAT declarations and large-sized firms show more auditing processes.

Back
Using the rate evasion/sales, firms that bunch at zero VAT declaration and small-sized firms evade more.
Results
Without LTD group

- Without LTD, group estimations rise 20pp approx.
- This shows we are obtaining a lower-bound.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VAT due</td>
<td>23.86</td>
<td>18.18</td>
<td>14.71</td>
<td>12.6</td>
<td>9.64</td>
<td>11.16</td>
<td>15.03</td>
</tr>
<tr>
<td>VAT</td>
<td>85.66</td>
<td>76.98</td>
<td>70.24</td>
<td>69.31</td>
<td>67.09</td>
<td>66.3</td>
<td>72.60</td>
</tr>
<tr>
<td>VAT proxy</td>
<td>90.65</td>
<td>74.13</td>
<td>67.05</td>
<td>70.37</td>
<td>66.35</td>
<td>64.07</td>
<td>72.10</td>
</tr>
<tr>
<td>Output VAT</td>
<td>38.56</td>
<td>28.39</td>
<td>24.79</td>
<td>23.56</td>
<td>19.7</td>
<td>18.61</td>
<td>25.60</td>
</tr>
</tbody>
</table>