Income inequality and taxation:
The case of South Africa

Janina Hundenborn*, Jon Jellema[†] , Ingrid Woolard*[§†‡]

*University of Cape Town, Southern Africa Labour and Development Research Unit (SALDRU),[†]Commitment to Equity Institute (CEQ), [§]IZA research fellow, [‡]UNU-WIDER

Research Questions

- How can data sets be combined to optimally analyse the trends in taxable incomes across the entire distribution?
- How do the assessment of taxable incomes differ between household surveys and tax administration data?

Background

South Africa has the highest income inequality in the world, even after progressive taxation and redistribution, the Gini coefficients remains highest compared to other middle income countries:

	Market income (1)	Net market income (2)	Disposable income (3)	Post-fiscal income (4)	Final income (5)
		(2)= (1) - Direct taxes	(3)= (2)+ Cash transfers	(4)= (3)- Indirect taxes	(5)=(4)+ in-kind transfers
Brazil (2009)	**0.579**	**0.565**	**0.544**	**0.546**	**0.439**
Costa Rica (2010)	0.508	0.500	0.489	0.486	0.393
El Salvador (2011)	0.440	0.436	0.430	0.429	0.404
Guatemala (2010)	0.551	0.550	0.546	0.551	0.523
Mexico (2010)	0.511	0.497	0.488	0.481	0.429
Peru (2009)	0.504	0.498	0.494	0.492	0.466
South Africa (2010)	**0.771**	**0.750**	**0.694**	**0.695**	**0.596**

Excerpt from Inchauste et al. (2015)

The South African government has substantially expanded fiscal programs and broadened the tax base to reduce poverty and inequality but these efforts have not translated into the equivalent results, as only a small fraction of the population gains directly from sustained economic growth and growth rates are lacking behind.

Measuring Inequality from Top to Bottom

Limited information available on the top end of the income distribution in household surveys due to lower response rates and lower incomes reported.

Tax data lacks information on individuals below the threshold but by default contains much more detail about filers at the top end.

Tax data estimates on the top tail can be combined with estimates on the bottom end obtained from household survey data in order to estimate the <u>entire</u> income distribution .

Diaz-Bazan (2015) defines the tax filing threshold as the optimal point to combine the two types of data rather than at different percentiles of the distribution. The Gini coefficient calculates as follows:

$$G[F] = \underbrace{\left(1 - F(b)\right)F(b)\left[\frac{\mu_2 - \mu_1}{\mu}\right]}_{\substack{\text{Adjustment term for the Gini based} \\ \text{on the two conditional distributions}}}$$

$$+ \underbrace{\left(F(b)^2 \frac{\mu_1}{\mu}\right)G[F_1]}_{\text{Gini in the household survey data}} + \underbrace{\left([1 - F(b)]^2 \frac{\mu_2}{\mu}\right)G[F_2]}_{\text{Gini in the tax administration data}}$$

Data Concerns

Household survey data captures individual taxable incomes relatively well except for the highest income groups.
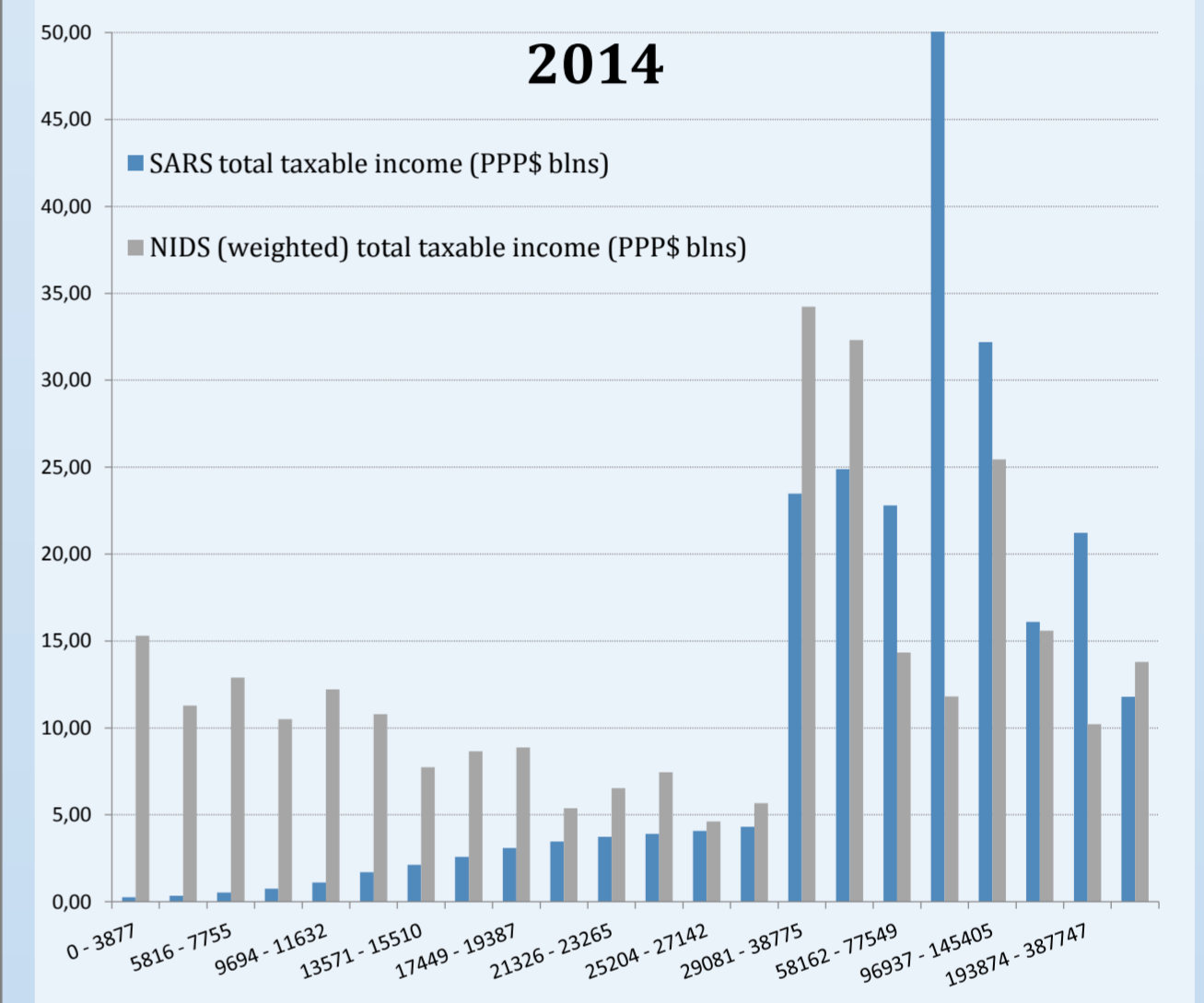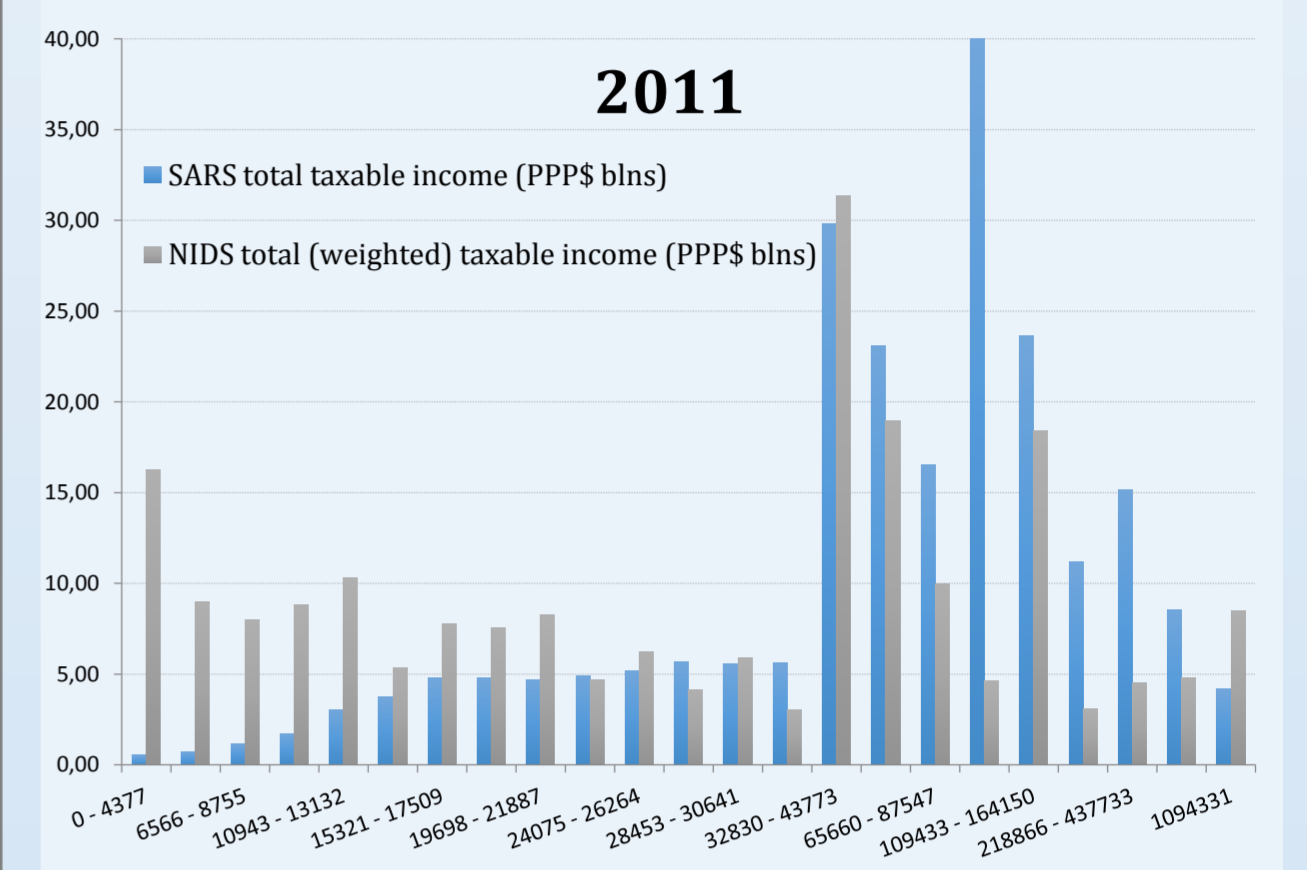
Source: Author's own calculations using NIDS W2 & W4, PIT 2011 & 2013, Tax Statistics 2015


- Re-weighted PIT data to match the 20% sample to the tax statistics, changes results significantly.

- Gini decreases between 2011 and 2014 but not if 2014 filing threshold is used on 2011 data.

- The 2014 filing threshhold is above the 99th percentile of the 2011 income distribution which increases the reliance on household survey data beyond optimal levels.

- Gini coefficient is rather large due to the fact that no social grants, remittances or other non-taxable transfers are included. A significant number of individuals at the bottom end of the income distribution report zero *taxable* income but may receive income from social grants or remittances. Any redistributive efforts by government are therefore not captured.

- Potential tax evasion may underestimate top incomes and thereby inequality. (Zucman et al., 2017).
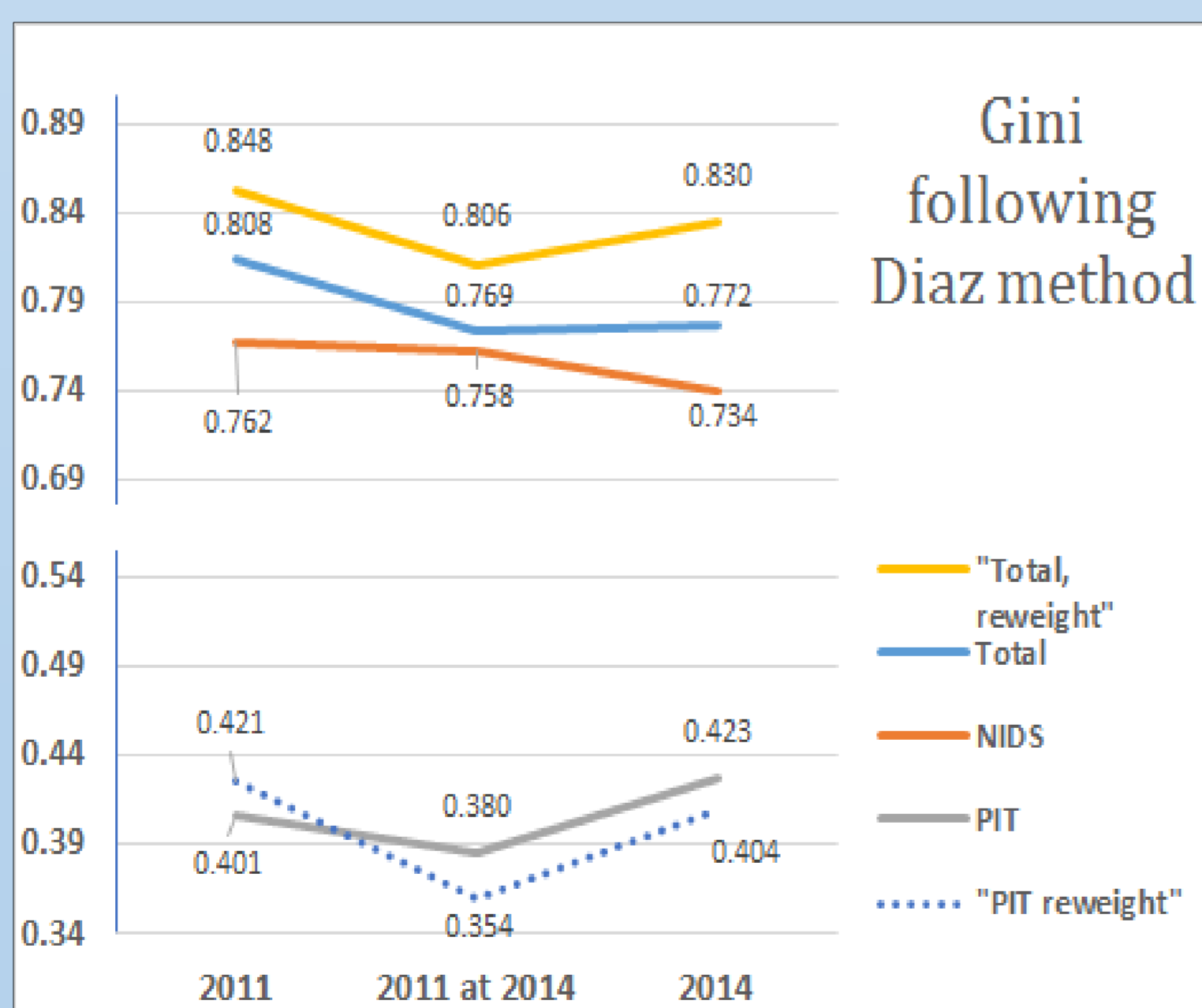
Conclusion and Policy Outlook

- Our analysis of the two types of data indicate that household survey data captures individual taxable income relatively well up until the higher income brackets where tax data appears to be more reliable.
- Absolute reduction in inequality between 2011 and 2014 when data sets are combined, however, the cause of this reduction is ambiguous. When 2011 data is assessed at the 2014 threshold, inequality is significantly lower.
- Gini coefficient of taxable income remains extremely high due to the exclusion of social grants, many report zero taxable income.
- Focus on income of individuals further excludes household resource sharing.
- The fact that inequality as measured by the Gini coefficient changes significantly once tax data is included in the analysis proves that the inclusion of their incomes is crucial in order to assess income inequality correctly.
- Analysis of taxes collected indicates the commitment to redistribution through progressive taxation.
- Progressive taxation can contribute significantly to a reduction in inequality combined with other measures such as progressive social spending. However, the possibility of expanding social spending is limited, especially given the small tax base and increasing concerns over tax payer compliance in an environment of potentially wasteful expenditure.

References
[1] DIAZ BAZAN, T. V. "Measuring inequality from top to bottom". World Bank Policy Research Working Paper, 2015.
[2] INCHAUSTE, G. AND LUSTIG, N. AND MABOSHE, M. AND PURFIELD, C. AND WOOLARD, I. "The Distributional Impact of Fiscal Policy in South Africa". World Bank Group, 2015.
[3] ZUCMAN, G. AND ALSTADSÆTER, A. AND JOHANNESEN , N. "Tax Evasion and Inequality", http://gabriel-zucman.eu/files/AJZ2017.pdf [09.06.2017], 2017

Data Sources
[1] SALDRU. National Income Dynamics Study 2010/2011, Wave 2 [dataset]. *Data First [distributor]*, Cape Town (2016). Version 3.1.
[2] SALDRU. National Income Dynamics Study 2014/2015, Wave 4 [dataset]. *Data First [distributor]*, Cape Town (2016). Version 1.1.
[3] SARS. Personal Income Tax (PIT) Data 2010 - 2011 [dataset]. *South African Revenue Service*, Pretoria (2016).
[4] SARS. Personal Income Tax (PIT) Data 2013 - 2014 [dataset]. *South African Revenue Service*, Pretoria (2016).