

South African Dataset for MAMS



AYODELE ODUSOLA
MARNA KEARNEY

SAM



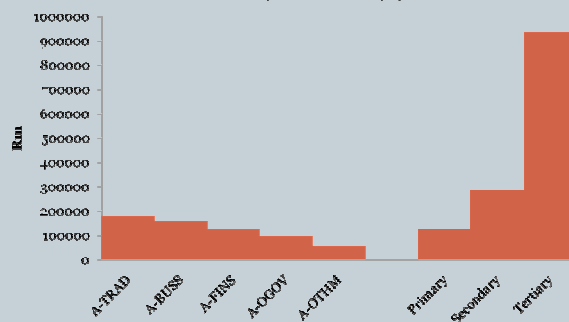
- Used 2005 Quantec SAM as base for MAMS SAM
- 46 Commodities and activities
- Government activities disaggregated
- Trade margins
- 4 Production factors (3 labour categories)
- Households, enterprises, government, RoW
- Factor quantities
- SAM is widely used by government agencies and academic institutions in SA
- Various data sources used in the compilation of SAM

Changes required for MAMS

- Developed new macroeconomic SAM to reflect changes required for MAMS
- Government accounts for education split into primary, secondary and tertiary
- Government account for water and sanitation – split water and sanitation into private and public provision
- Government infrastructure
- Education split from personal services in SAM – then divided into primary, secondary and tertiary
- Private health split from medical and other services industry in SAM
- Savings and investment block added
 - Allocation of GDFI across institutions responsible
- Interest payment block added
- Aggregation of commodities and activities in SAM

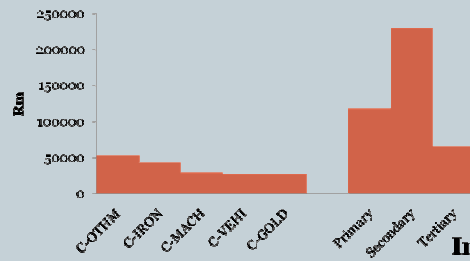
SAM in Pictures

Value Added by Industry / Sector

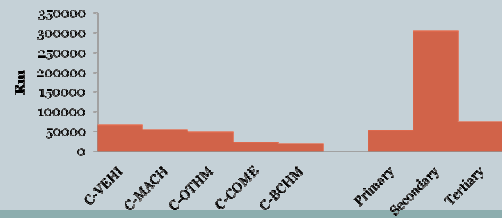


SAM in Pictures

Exports by Commodity / Sector

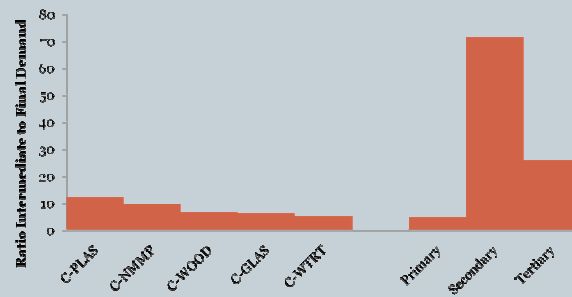


Imports by Commodity / Sector



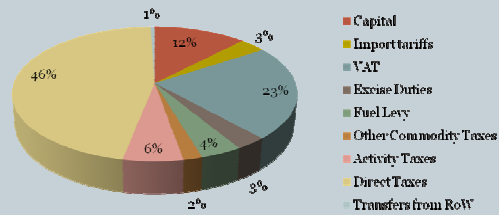
SAM in Pictures

Intermediate to Final Demand by Commodity / Sector

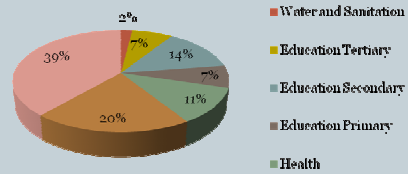


SAM in Pictures

Sources of Revenue for Government

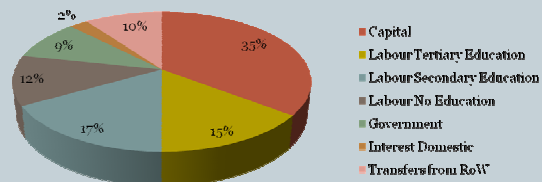


Value Added by Government Function

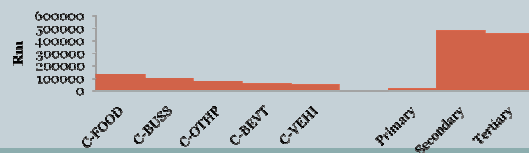


SAM in Pictures

Sources of Income for Households

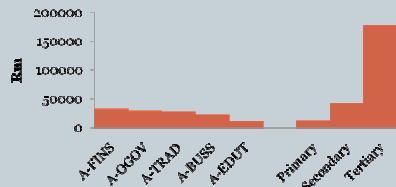


Household Consumption by Commodity / Sector

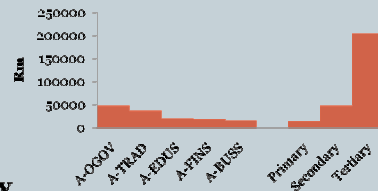


SAM in Pictures

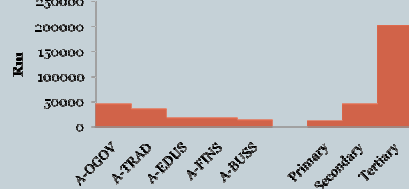
Remuneration Tertiary Education



Remuneration Secondary Education

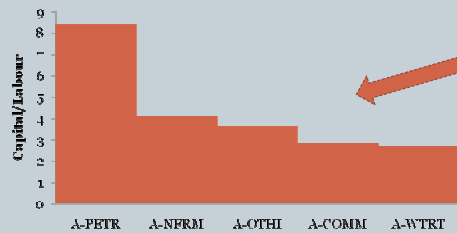


Remuneration Primary Education

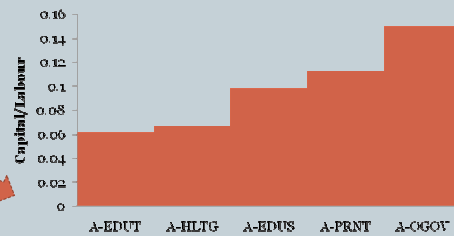


SAM in Pictures

Capital Labour Ratio



Capital Labour Ratio



SAM in Pictures

Savings and Investment Block

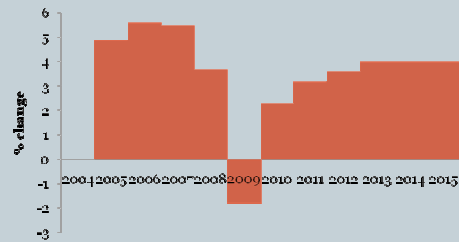
	HHH	GOV	INT-DOH	INT-ROW	SAV-HHH	SAV-GOV	SAV-ROW	CAP-HHH	CAP-GOV	CAP-ROW	INV-PRV	INV-HHS	INV-EDUH	INV-EDUS	INV-EDUP	INV-HLTG	INV-CHNF	INV-OGG	ROW	TOTAL
C											193233	4221	1409	3081	4006	4125	29463	17052		296590
HHH			37417.21																	37417
GOV																				0
INT-DOH			37417																	37417
INT-ROW			156956	13536																170534
SAV-HHH			219187																	219187
SAV-GOV			-13332																	-13332
SAV-ROW																			64370	64370
CAP-HHH					219187						8181									227378
CAP-GOV					-13332			60045		19593										62307
CAP-ROW							64370													64370
INV-PRV								152647		40586										193233
INV-HHS											4221									4221
INV-EDUH											1409									1409
INV-EDUS											3081									3081
INV-EDUP											4006									4006
INV-HLTG											4125									4125
INV-CHNF											29463									29463
INV-OGG											17052									17052
DSTK										14686	-1051									13635
ROW																				170534
TOTAL	376184	37623	37417	170534	219187	-13332	64370	227378	62307	64370	193233	4221	1409	3081	4006	4125	29463	17052	64370	170534

General Dataset

- SAM balanced using entropy
- GDP growth forecast based on forecast by National Treasury published in February 2010
- Population and Labour Growth Forecast based on data and model from Actuarial Society of South Africa
- Elasticities from existing sources – National Treasury CGE model, Thurlow (2003), De Wet (2003), K Gibson (2001)
- Labour Quantities provided by Quantec with SAM

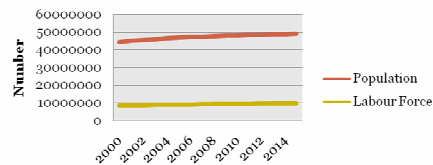
General Dataset in Pictures

GDP at market prices (real)

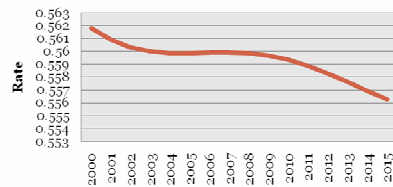


General Dataset in Pictures

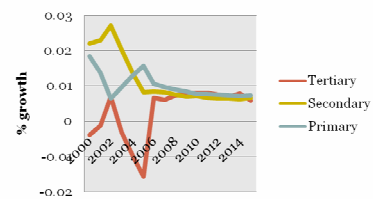
Population and Labour Force Forecast



Labour Force Participation Rate

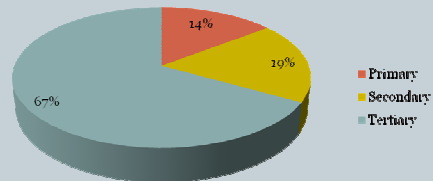


Labour Force Growth Rate

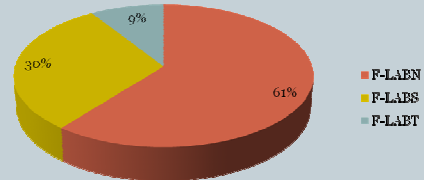


General Dataset in Pictures

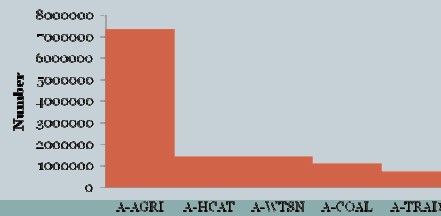
Employment Share by Sector



Employment Share by Education Level

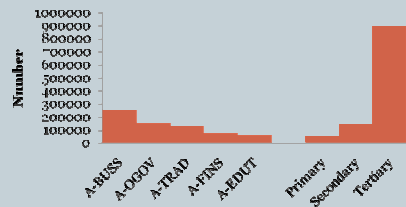


Total Employment by Industry

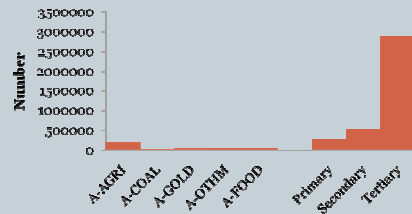


General Dataset in Pictures

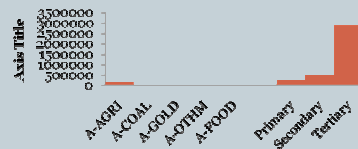
Employment Tertiary Education



Employment Secondary Education



Employment Primary Education



MDG Dataset

- Various data sources used
 - Education publications – 2000 – 2007
 - General household surveys 2008
 - UN datasets
 - Statistics South Africa population time series
 - SARB
- Child mortality estimation – household survey does not cover this

MDG Dataset Estimations - mdgeduelas

```
. logit wateraccess Income Prov Popgrp GovspendwaterSanitationc_in electriccityaccess

Iteration 0: log likelihood = -17149.577
Iteration 1: log likelihood = -16349.555
Iteration 2: log likelihood = -16142.598
Iteration 3: log likelihood = -15994.331
Iteration 4: log likelihood = -15862.071
Iteration 5: log likelihood = -15861.885
Iteration 6: log likelihood = -15861.885

Logistic regression               Number of obs   = 64913
                                Lk chi2(1)       = 6375.18
                                Prob > chi2       = 0.0000
                                Pseudo R2        = 0.1917

Log Likelihood = -15861.885

wateraccess      Conf. Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
Income          -.0000274   1.68e-06   16.31   0.000   -.0000241   -.0000307
Prov            -2.740775   .0607064   25.40   0.000   -2.862225   -2.617226
Popgrp          2.001304   .0954864   20.92   0.000   1.813841    2.188768
Govspendwaterc  .1053803   .0134239   8.09    0.000   .08227      .128496
electricity     1.903997   .0325078   58.41   0.000   1.840107    1.967888
_cons          -2.92108    .1654051   -17.67   0.000   -3.247423   -2.594802

Note: 0 failures and 79 successes completely determined.

. margins, dydx(*)
      Average marginal effects      Number of obs   = 64913
      Model VCE : OIM

Expression : Pr(wateraccess), predict()
dy/dx w.r.t. : Income Prov Popgrp GovspendwaterSanitationc_in electriccityaccess

      _delta-method_
      dy/dx      Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
Income          1.68e-06   1.00e-07   16.22   0.000   1.68e-06   1.68e-06
Prov            .0333805   .0046039   25.77   0.000   .0241765   .0425845
Popgrp          .1190704   .0057564   20.68   0.000   .107768    .1303727
Govspendwaterc  .0646621   .0070797   9.10    0.000   .0504827   .0788416
electricity     .111281    .0019304   58.08   0.000   .1084562   .1141057

. margins, eyex(*)
      Average marginal effects      Number of obs   = 64913
      Model VCE : OIM

Expression : Pr(wateraccess), predict()
ey/ex w.r.t. : Income Prov Popgrp GovspendwaterSanitationc_in electriccityaccess

      _delta-method_
      ey/ex      Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
Income          .0062322   .0001594   39.23   0.000   .0059398   .0065646
Prov            .0578219   .0030053   28.88   0.000   .0519836   .0636599
Popgrp          .1533847   .0072171   21.06   0.000   .1394055   .1673639
Govspendwaterc  .06639    .0081535   8.17    0.000   .0506095   .0821705
electricity     .0648014   .0009386   69.42   0.000   .0629354   .0666674
```

onpc_in elect

```
Iteration 0: log likelihood = -15571.225
Iteration 1: log likelihood = -15329.985
Iteration 2: log likelihood = -12958.637
Iteration 3: log likelihood = -12858.239
Iteration 4: log likelihood = -12850.888
Iteration 5: log likelihood = -12850.861
Iteration 6: log likelihood = -12850.861
```

sanitation-s	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
Income	.0000145	1.30e-06	11.16	0.000	.0000119 .0000170
Proc	.0001327	.000706	0.13	0.895	.0002435 -.000713
Popprp	0.139477	.0654568	16.08	0.000	.0119322 0.1650212
Gvpspendln	.2591146	.0134212	19.31	0.000	.2328096 .2854197
electricit-s	0.209878	.0346984	60.52	0.000	0.138187 0.2678885
_cons	-2.625682	.1349623	-17.52	0.000	-2.934946 -2.331917

```
. margins, dydx(*)
```

```
Expression : Pr(sanitationaccess), predict()
dy/dx w.r.t. : Income Prov Popgrp Govspendwatersanitationc_in electricityacce
```

```

. margins, ayex(*)
Average marginal effects
Model VCE      : OIM
Number of obs   =    64913
Expression      : Pr(sanitationaccess). predict()

```

	ey/ex	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
Totals	0010702	0002011	20.04	0.000	0016722	0014683

```
end of do-file
```



```
Number of STRATA = 1                      Number of obs = 40693
```

Prob > F = 0.0000

```

- margins, dydx(*)
Average marginal effects
Number of obs = 4069
Model VCE = linearized
Expression = Pr(attendprimaryschool), predict()
dy/dx w.r.t. = Gender q124dica Prov Popgrsp headattendeedschool spouseattendeedschool

```

```

. margins, at(x*)
      Average marginal effects
      model VCE      : linearized
      number of obs   =    40692

Expression : Pr(attendprimaryschool), predict()
dy/dx w.r.t. : Gender q124dica Prov Popgrq headattendeedschool spouseattendeedschool

```


 University of Twente
 Enschede, The Netherlands

MDG Dataset Estimations - mdgeduelas

```
svy: logit attendsecondaryschool cgender q23afia vrow popgrp headattendedschool spouseattendedschool edu_qual pc_income_in infrastructureaccess wp_3_b  
(running logit on estimation sample)
```

```
Survey: logistic regression  
Number of strata = 1  
Number of obs = 2912  
Population size = 2236181  
Design df = 2911  
P < 0.05 = 2902  
P > 0.10 = 48.18  
P > 0.001 = 0.000
```

	Coef.	linearized std. err.	z	P> z	[95% Conf. Interval]
gender	-.1909807	.0373007	-4.89	0.000	-.2649273 - .1170961
q23afia	.1274811	.0378843	3.36	0.000	.0530882 .2018740
vrow	-.0323941	.0336037	-0.96	0.347	-.0988009 .0347288
popgrp	.0110411	.0327304	0.34	0.690	-.0539444 .0759610
headattend-1	.0510143	.0378458	1.35	0.181	-.0237771 .1258417
spouseatt-1	.0717028	.0374549	1.91	0.056	-.0337027 .1771083
edu_qual	-.0440005	.0366417	-1.20	0.230	-.1168182 .0287771
pc_income_in	-.1440005	.0366417	-3.93	0.000	-.2158182 -.0721818
infrastructure	.0311111	.0359445	0.87	0.385	-.0393861 .1015977
wp_3_b	.0717028	.0374549	1.91	0.056	-.0337027 .1771083
_cons	-1.718838	.2403966	-6.88	0.000	-2.201276 -1.235993

```
. margins, dydx(*)  
Average marginal effects  
Model VCE : linearized  
Expression : pr(attendsecondaryschool), predict()
```

```
dy/dx w.r.t. : cgender q23afia vrow popgrp headattendedschool spouseattendedschool edu_qual pc_income_in infrastructureaccess wp_3_b
```

		delta-method				
	dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]	
gender	-.0184456	.0038617	-4.78	0.000	-.0260244	-.0108708
q23afia	.0078111	.0038617	2.02	0.044	.0001841	.0154381
vrow	.0032121	.0038617	0.83	0.408	-.0034671	.0098913
popgrp	.0002041	.0038617	0.05	0.959	-.0074671	.0078753
headattend-1	.0034671	.0038617	0.90	0.369	-.0034671	.0074671
spouseatt-1	.0040000	.0038617	1.04	0.300	-.0034671	.0114671
edu_qual	-.0034671	.0038617	-0.90	0.369	-.0074671	.0005329
pc_income_in	-.0034671	.0038617	-0.90	0.369	-.0074671	.0005329
infrastructure	.0034671	.0038617	0.90	0.369	.0005329	.0064329
wp_3_b	.0073121	.0049139	1.51	0.127	-.0021564	.0170204

```
. margins, eyex(*)  
Average marginal effects  
Model VCE : linearized  
Expression : pr(attendsecondaryschool), predict()
```

```
ey/ex w.r.t. : cgender q23afia vrow popgrp headattendedschool spouseattendedschool edu_qual pc_income_in infrastructureaccess wp_3_b
```

	delta-method					
	ey/ex	std. err.	z	p> z	[95% Conf. Interval]	
gender	-.2406494	.0503019	-4.78	0.000	-.3382193	-.1420595
q23afia	.0471718	.0503019	0.94	0.345	-.0511861	.1455297
vrow	.0124261	.0503019	0.25	0.800	-.0927722	.1176244
popgrp	.0037041	.0503019	0.07	0.944	-.0927722	.1001804
headattend-1	.0040000	.0503019	0.08	0.938	-.0927722	.1007722
spouseatt-1	.0040000	.0503019	0.08	0.938	-.0927722	.1007722
edu_qual	-.0040000	.0503019	-0.08	0.938	-.1007722	.0927722
pc_income_in	-.0040000	.0503019	-0.08	0.938	-.1007722	.0927722
infrastructure	.0040000	.0503019	0.08	0.938	.0927722	.1007722
wp_3_b	.0040000	.0503019	0.08	0.938	.0927722	.1007722

```
.  
.
```

MDG Dataset Estimations - mdgeduelas

```
svy: logit attendsecondary cgender q23afia vrow popgrp headattendedschool spouseattendedschool edu_qual pc_income_in infrastructureaccess wp_uni_b  
(running logit on estimation sample)
```

```
Survey: logistic regression  
Number of strata = 1  
Number of obs = 2912  
Population size = 2236181  
Design df = 2911  
P < 0.05 = 2902  
P > 0.10 = 48.18  
P > 0.001 = 0.000
```

attendedu-y	Coef.	Linearized Std. Err.	z	P> z	[95% Conf. Interval]
gender	.1886473	.0461322	2.82	0.005	.0969371 .2803575
q23afia	.1146079	.0461322	2.48	0.000	.0224361 .2067797
vrow	.0449911	.0461322	0.98	0.328	-.0449911 .1349733
popgrp	-.1486611	.0461322	-3.21	0.001	-.2392311 -.0624511
headattend-1	.0702811	.0461322	1.49	0.143	-.0218936 .1684556
spouseatt-1	.0702811	.0461322	1.49	0.143	-.0218936 .1684556
edu_qual	-.0702811	.0461322	-1.52	0.263	-.1419204 .0013582
pc_income_in	.0702811	.0461322	1.52	0.060	.0013582 .1392040
infrastructure	.0702811	.0461322	1.52	0.060	.0013582 .1392040
wp_uni_b	.0702811	.0461322	1.52	0.060	.0013582 .1392040
_cons	-4.433642	.4881774	-9.09	0.000	-5.400000 -3.467284

```
. margins, dydx(*)  
Average marginal effects  
Model VCE : linearized  
Expression : pr(attendsecondary), predict()
```

```
dy/dx w.r.t. : cgender q23afia vrow popgrp headattendedschool spouseattendedschool edu_qual pc_income_in infrastructureaccess wp_uni_b
```

dy/dx u.c.r.t. : Gender q23afia row popgrp headattendchool spouseattendchool					
	dy/dx	std. err.	z	P> z	[95% Conf. Interval]
gender	.0083116	.0042423	1.91	0.055	.000207 .016412
q23afia	.0033333	.0042423	0.81	0.418	-.007742 .014409
row	.0033333	.0042423	0.79	0.430	-.007742 .014409
popgrp	-.0051429	.0042423	-1.20	0.227	-.013544 .003258
headattend-1	.0026667	.0042423	0.63	0.524	-.008451 .013784
spouseatt-1	.0026667	.0042423	0.63	0.524	-.008451 .013784
edu_qual	-.0055633	.0049566	-1.12	0.262	-.015279 .004153
pc_income_in	.0033333	.0042423	0.79	0.430	-.007742 .014409
infrastructure	.0022491	.0051101	0.44	0.658	-.011484 .016002

```
. margins, eyex(*)  
Average marginal effects  
Model VCE : linearized  
Expression : pr(attendsecondary), predict()
```

```
ey/ex w.r.t. : cgender q23afia vrow popgrp headattendedschool spouseattendedschool edu_qual pc_income_in infrastructureaccess wp_uni_b
```

	ey/ex	std. err.	z	p> z	[95% conf. interval]
gender	.2889739	.095179	3.04	0.002	.1000000 .4779478
q23afia	.0479701	.121746	0.39	0.695	-.1978881 .2928282
vrow	.1208089	.095179	1.27	0.204	-.0540000 .2958089
popgrp	.0702811	.095179	0.74	0.459	-.1000000 .2405621
headattend-1	.0702811	.095179	0.74	0.459	-.1000000 .2405621
spouseatt-1	.0702811	.095179	0.74	0.459	-.1000000 .2405621
edu_qual	-.0702811	.095179	-0.74	0.459	-.2405621 .1000000
pc_income_in	.0702811	.095179	0.74	0.459	.1000000 .2405621
infrastructure	.0702811	.095179	0.74	0.459	.1000000 .2405621
wp_uni_b	.0702811	.095179	0.74	0.459	.1000000 .2405621

Microsimulation



- Previous micro simulation models were developed for SA based on 2000 data – used 2000 IES combined with 2000 LFS
- Newer dataset is available IES 2005/2006, but concerns as this survey does not sufficiently cover employment variables and is not linked with LFS

Conclusion



- Need to finalise estimations for mdg dataset
- Test data in CGE model
- Data verification with Departments etc.
- Updated as more /better sources of data becomes available
- Capacity building